

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

1 nucleic - nucleic search, using sw model  
in on: February 1, 2004, 19:35:31 ; Search time 12453 Seconds  
(without alignments)  
11484.787 Million cell updates/sec  
tle: US-09-042-460-1  
arfect score: 3496  
quence: 1 GAATTCGGGTGGAGGCC.....CCGAGCTCGTACCAAGCTT 3496  
oring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
sarched: 2888711 seqs, 2045481386 residues  
otal number of hits satisfying chosen parameters: 5777422  
inimum DB seq length: 0  
aximum DB seq length: 2000000000  
ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.on.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	3426	98.0	3426	10	AF051911	Mus muscu	
2	3367.4	96.3	3369	10	AF073311	Mus muscu	
3	2578.6	73.8	4170	10	AF149012	Mesocric	
4	1758	50.2	1758	10	AF247818	Rattus no	
5	1603.6	45.9	8960	12	AF043739	Synthetic	
6	1600.8	45.8	4015	6	AR104587	Sequence	
7	1600.8	45.8	4015	6	AR175848	Sequence	
8	1600.8	45.8	4015	6	AR182221	Sequence	
9	1600.8	45.8	4015	6	AR224455	Sequence	
10	1600.8	45.8	4015	6	AR226390	Sequence	
11	1600.8	45.8	4015	6	AR243328	Sequence	
12	1600.8	45.8	4015	6	AR263555	Sequence	
13	1600.8	45.8	4015	6	AR265996	Sequence	
14	1600.8	45.8	4015	6	AX019310	Sequence	
15	1600.8	45.8	4015	6	AX133979	Sequence	
16	1600.8	45.8	4015	6	AX552695	Sequence	
17	1600.8	45.8	4015	6	BD011044	Human tel	
18	1600.8	45.8	4015	6	BD015832	Quantitat	
19	1600.8	45.8	4015	6	BD082985	Method fo	
20	1600.8	45.8	4015	6	BD131727	Method fo	
21	1600.8	45.8	4015	6	E36793	Human telom	
22	1600.8	45.8	4015	9	AF015950	Homo sapi	
23	1600.8	45.8	4027	9	AF018167	Homo sapi	
24	1600.8	45.8	4042	6	AX001446	Sequence	
25	1600.8	45.8	4042	6	AX003121	Sequence	
26	1600.8	45.8	4042	6	BD136185	Human tel	
27	1600.8	45.8	4070	6	AX391846	Sequence	
28	1596	45.7	3399	6	AX481414	Sequence	
29	1596	45.7	4037	6	BD011070	Human tel	
30	1594.6	45.7	4037	6	E36819	Human telom	
31	1594.6	45.6	3396	6	BD091553	Adult bon	
32	1594.6	45.6	3396	6	BD094749	The cell	
33	1594.6	45.6	3396	6	BD096291	Cells cap	
34	1593.6	45.6	13766	6	AX553919	Sequence	
35	1584.8	45.3	8742	6	AX553924	Sequence	
36	1520.2	43.5	4029	6	AR104586	Sequence	
37	1520.2	43.5	4029	6	AR175847	Sequence	
38	1520.2	43.5	4029	6	BD011069	Human tel	
39	1520.2	43.5	4029	6	E36818	Human telom	
40	1462.6	41.8	7498	6	BD123426	Transgeni	
41	1462.6	41.8	7498	10	AF121949	Mus muscu	
42	1462.6	41.8	216363	2	AC131076	Mus muscu	
43	1323.2	37.8	3855	6	AR243330	Sequence	
44	1323.2	37.8	3855	6	BD011046	Human tel	
45	1323.2	37.8	3855	6	E36795	Human telom	

ALIGNMENTS

RESULT 1  
AF051911  
LOCUS AF051911 3426 bp mRNA linear ROD 01-DEC-2000  
DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.  
ACCESSION AF051911  
VERSION AF051911.1 GI:3005591  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 3426)  
REFERENCE  
AUTHORS Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and DePinho,R.A.  
TITLE Expression of mouse telomerase reverse transcriptase during  
development, differentiation and proliferation

JOURNAL  
 3DLINE  
 PUBMED  
 9582020  
 ERENCE  
 2 (bases 1 to 3426)  
 GREENBERG, R.A., ALLSOPP, R.C., CHIN, L., MORIN, G.B. and DEPINHO, R.A.  
 Direct Submission  
 Submitted (02-MAR-1999) Microbiology and Immunology, Albert  
 Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY  
 10461, USA

TURES  
 source  
 1. 3426  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="13"  
 /map="Mdb offset position 40; 40 cm from centromere"  
 30. 3398  
 /note="mPRT"  
 /codon\_start=1  
 /product="telomerase reverse transcriptase"  
 /protein\_id="AAC09323.1"  
 /db\_xref="GI:300592"  
 /translation="MTRAPCPVAVSLRSLRYSRVWPLATFVRLGPEGRRLVQGPDP  
 KYETVAOCLVMHWSQPPADLSFHOSSSLKELVARVVOPLCRERNVLAQFE  
 LNAEAGSPMAFTSVRSYLPNVIVETLAVSGAMLLLSRVGDDLLVLLHCLAYL  
 LVFPCAYQCCSPFIQCATIDWPSVASIRFTRVGRNFTNLRPLQIIXSSRSP  
 ARPEVPTAKDLSSKGVSLDSLSGSCVCKKPSSTLSLSPRQVLPFTPGKSWPSP  
 FLYSRGQERLNPFLSNLQPNLTGAARLVEIIFLGRSPRTSGPLCHRLSRYW  
 QMPLFQQLLVNHAQVYVLRSHCHRPRTANQQVTDALNTSPHMLDLRLHSPWQ  
 VYGLFACLCVKVYASLGRTHRRFRFPLKKEFISLKYKLSLOELMWMKVEDCH  
 WLRSSPKDRVPAARHLRERILATFLMLDVTYVOLLRSFVIFESTFQKNRUFY  
 RKSVMKLSIGVROHLRVLRLSLOEVARHODTWLAMPICRLRFLPKNGRLPIV  
 NMSYNGTRALGRKQAFQRLKTLFSLMYNTRKHPLHSGSVLGNMDLYRTWA  
 FVLVRALDQTRMYEVKADVTGAVDALPQGLKLEWVWAMIRSESTYCIQYAVR  
 DSQGVHFSRQVITLSDLPYMGQFLKLOSDASALANSVVEIOSISWNESSSL  
 PDFFLFSLHRSVYKCTCCQIPQGSLSLILCSLFCGDMENKSFVQRDGLL  
 LRVDFFLVLTHLOQATFLSTLVHGVPEYGMINLQNTVNVFVPEPTLGGAPYQ  
 LPAHCLFPCWGLDITLQVFCVSYQAQTSIKTSIQSVFPAKTRNKLKSLVR  
 LKCHGLFLDLQVNSLQVCIYKIFLQAYRFHACVILQPDFQVRKNTFFFLGIS  
 SQACCYALIKVKNPMTKASGSPPEAAHWCYQAFLLKLAHSVIYKCLGLPLRT  
 AQLLCRKLPEATMILKAAADPALSLDFQILD"

ASE COUNT 746 a 979 c 907 g 794 t  
 UGIN

Query Match 98.0%; Score 3426; DB 10; Length 3426;  
 Best Local Similarity 100.0%; Fred. No. 0;  
 Matches 3426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 10 GTGGAGGCCATCCCGGCTTTGAGCAATGACCGCGCTCTCTGTTGCGCGGCTGC 69  
 b 1 GTGGAGGCCATCCCGGCTTTGAGCAATGACCGCGCTCTCTGTTGCGCGGCTGC 60  
 Y 70 GCTCTCTGCTGCGAGCCGATACCGGAGGTGTGCGCGCTGCGCACTTTGTCGCGGCG 129  
 b 61 GCTCTCTGCTGCGAGCCGATACCGGAGGTGTGCGCGCTGCGCACTTTGTCGCGGCG 120  
 Y 130 TGGGGCCGAGGCGCGCTTTGCAACCGGGGACCCGAGATCTACCGCACTTTGG 189  
 b 121 TGGGGCCGAGGCGCGCTTTGCAACCGGGGACCCGAGATCTACCGCACTTTGG 180  
 Y 190 TTGCCCAATGCTAGTGTGATGCACTGGGGCTCACAGCTTCCAGCTTCCCTTCTCT 249  
 b 181 TTGCCCAATGCTAGTGTGATGCACTGGGGCTCACAGCTTCCAGCTTCCCTTCTCT 240  
 Y 250 TCCACCAAGGTGTCTATCCCTGAAGAGCTGTGGCGAGGTTGCGAGACTTGGCGAGC 309  
 b 241 TCCACCAAGGTGTCTATCCCTGAAGAGCTGTGGCGAGGTTGCGAGACTTGGCGAGC 300  
 Y 310 GCAACGAGAGAAACGCTGCTTTGGCTTTGAGCTGCTTTAACAGGCCAGAGCGCGGC 369  
 b 301 GCAACGAGAGAAACGCTGCTTTGGCTTTGAGCTGCTTTAACAGGCCAGAGCGCGGC 360

QY 370 CTCCTCATGGCCTTCACTAGTAGCGGTAGCTACTTGGCCAAACACTGTTATTGAGACCC 429  
 Db 361 CTCCTCATGGCCTTCACTAGTAGCGGTAGCTACTTGGCCAAACACTGTTATTGAGACCC 420  
 QY 430 TCGGTGTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 489  
 Db 421 TCGGTGTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 480  
 QY 490 ACTCTGTGCGACACTGTGCTTATCTTCTGTGTGCGCCCGCCAGCTGTGCGTACCAAGTGT 549  
 Db 481 ACTCTGTGCGACACTGTGCTTATCTTCTGTGTGCGCCCGCCAGCTGTGCGTACCAAGTGT 540  
 QY 550 GTGGGTCTCCCTGTGTACCAATTTGTGCGCACACAGGATATCTGGCCCTCTGTGTGCGCTA 609  
 Db 541 GTGGGTCTCCCTGTGTACCAATTTGTGCGCACACAGGATATCTGGCCCTCTGTGTGCGCTA 600  
 QY 610 GTTACAGCCCGACCCGACCCGTCGAGGAGATTTTCACTAACCTTAGTGTCTTACACAGA 669  
 Db 601 GTTACAGCCCGACCCGACCCGTCGAGGAGATTTTCACTAACCTTAGTGTCTTACACAGA 660  
 QY 670 TCAAGAGCAGTAGTGTGCGCAGGAGACCGGACCCCTGCGCTTGCCTATCTCGAGGTACAA 729  
 Db 661 TCAAGAGCAGTAGTGTGCGCAGGAGACCGGACCCCTGCGCTTGCCTATCTCGAGGTACAA 720  
 QY 730 AGAGGCATCTGAGTGTCTCACAGTACAGTGTGCTTACAGTAAAGAGCCAGATGCTATC 789  
 Db 721 AGAGGCATCTGAGTGTCTCACAGTACAGTGTGCTTACAGTAAAGAGCCAGATGCTATC 780  
 QY 790 CTGTCTCCGAGAGTGTGAGGAGGAGACCCCAAGGAGGAGTGTCTACCAACCCCATCAGGCAAA 849  
 Db 781 CTGTCTCCGAGAGTGTGAGGAGGAGACCCCAAGGAGGAGTGTCTACCAACCCCATCAGGCAAA 840  
 QY 850 CATGGGTGCCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909  
 Db 841 CATGGGTGCCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 910 TTAAGAGAAAGTGTGTGACCTGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969  
 Db 901 TTAAGAGAAAGTGTGTGACCTGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 970 GCTTCAATCTCTGCTGTGTGACCAACCCCGGCAAAATGCTTTCAGCTCAGGCAATTTATTG 1029  
 Db 961 GCTTCAATCTCTGCTGTGTGACCAACCCCGGCAAAATGCTTTCAGCTCAGGCAATTTATTG 1020  
 QY 1030 AGACAGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1089  
 Db 1021 AGACAGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
 QY 1090 TACTCAGCAACCTTCCAGCCTTAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
 Db 1081 TACTCAGCAACCTTCCAGCCTTAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 QY 1150 TGGGTCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCGTCTATCGCTGCTGAT 1209  
 Db 1141 TGGGTCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCGTCTATCGCTGCTGAT 1200  
 QY 1210 ACTGGCAGATGCGGCGCTTGTTCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269  
 Db 1201 ACTGGCAGATGCGGCGCTTGTTCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1270 TCAGACTCTCTCAGGTCTACATTTGAGGTTTCGAAACAGCAACCAACAGGTGACAGATGCTCT 1329  
 Db 1261 TCAGACTCTCTCAGGTCTACATTTGAGGTTTCGAAACAGCAACCAACAGGTGACAGATGCTCT 1320  
 QY 1330 TGAACACGAGCCGACCGACCTCATGAGATTTCTCGGCTGCAAGAGGCTTCTTGGGCTGAG 1389  
 Db 1321 TGAACACGAGCCGACCGACCTCATGAGATTTCTCGGCTGCAAGAGGCTTCTTGGGCTGAG 1380  
 QY 1390 TATATGTTTCTTGGGCTGCTGCTCTGCAAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
 Db 1381 TATATGTTTCTTGGGCTGCTGCTCTGCAAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 QY 1450 GGCACATGAGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1509



[illegible]

	2410	2401	2470	2461	2530
y	TGAATGAGAGCAGCAGCGCTGTTGACTTCTTCTGCACTTCTCTGCGTACAGTGTGC				
b					
	2469	2460	2529	2520	2589
y	TGAATGAGAGCAGCAGCGCTGTTGACTTCTTCTGCACTTCTCTGCGTACAGTGTGC		TAAAGATTGCTGACAGTGCTATACGAGTGCCAGGCGTCCCGCAGGCTTAT		
b					
	2470	2461	2530	2589	
y	TGAATGAGAGCAGCAGCGCTGTTGACTTCTTCTGCACTTCTCTGCGTACAGTGTGC		TAAAGATTGCTGACAGTGCTATACGAGTGCCAGGCGTCCCGCAGGCTTAT	CCACCGCTCTGAGTCTGTGTTTCGGACATGGAGAACAGCTTTTCTGAGGTGC	
b					

RESULT 2

AF073311

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AF073311

Mus musculus telomerase catalytic subunit mRNA, complete cds.

AF073311

AF073311.1

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

3369 bp

linear

ROD 09-SEP-1998



1479 TTAAGAGGTTCTATCTCGTTGGGAAATACGCAAGCTATCACTCGAGAACTGATGCG 1538  
1441 TTAAGAGGTTCTATCTCGTTGGGAAATACGCAAGCTATCACTCGAGAACTGATGCG 1500  
1539 AAGATGAAGTACGAGGATGCGCACTGGCTCCGACAGAGCCCGGGAAGACCGTGTCCCG 1598  
1501 AAGATGAAGTACGAGGATGCGCACTGGCTCCGACAGAGCCCGGGAAGGACCGTGTCCCG 1560  
1599 GCTGCAGAGCACCGTCTGAGGGAGAGGATCTGTGGTACGTTCTCTGTCTGTGCTGATGAC 1658  
1561 GCTGCAGAGCACCGTCTGAGGGAGAGGATCTGTGGTACGTTCTCTGTCTGTGCTGATGAC 1620  
1659 ACATACGCTGACAGCTGTGTTAGTCACTCTTTTATCATCACAGAGGACATCCAGAG 1718  
1621 ACATACGCTGACAGCTGTGTTAGTCACTCTTTTATCATCACAGAGGACATCCAGAG 1680  
1719 AACAGGCTCTTTCTTACCGTAAGAGTGTGTGAGCAAGCTGCAGAGCATTTGGAGTCAAG 1778  
1681 AACAGGCTCTTTCTTACCGTAAGAGTGTGTGAGCAAGCTGCAGAGCATTTGGAGTCAAG 1740  
1779 CAACACCTTGAGAGTGTGGCTTACCGGAGCTGTCAAGAGAGAGTCAAGCATCACAG 1838  
1741 CAACACCTTGAGAGTGTGGCTTACCGGAGCTGTCAAGAGAGAGTCAAGCATCACAG 1800  
1839 GACACCTGGCTAGCATTGCCATCTGCAGACTGGCTTATCCCAAGCCCAAGCCGCTG 1898  
1801 GACACCTGGCTAGCATTGCCATCTGCAGACTGGCTTATCCCAAGCCCAAGCCGCTG 1860  
1899 CGGCCCATTTGAACTGAGTATAGCATGGGTACAGAGCTTTTGGGCGAGAGGAGCAG 1958  
1861 CGGCCCATTTGAACTGAGTATAGCATGGGTACAGAGCTTTTGGGCGAGAGGAGCAG 1920  
1959 GCCAGCATTTCCCGCAGCTTCAAGACTCTCTTCAGCATGCTCAACTATGAGCGGACA 2018  
1921 GCCAGCATTTCCCGCAGCTTCAAGACTCTCTTCAGCATGCTCAACTATGAGCGGACA 1980  
2019 AAACATCTCACCTTATGGGGTCTTCTGTACTGGGTATGAATGACATCTACAGGACCTGG 2078  
1981 AAACATCTCACCTTATGGGGTCTTCTGTACTGGGTATGAATGACATCTACAGGACCTGG 2040  
2079 CGGGCTTTGTGCTGCTGCTGCTGCTGAGACAGACACCCAGGATGACTTTGTTAG 2138  
2041 CGGGCTTTGTGCTGCTGCTGCTGAGACAGACACCCAGGATGACTTTGTTAG 2100  
2139 GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGTAAGCTGTGGAGTGTGTTGCC 2198  
2101 GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGTAAGCTGTGGAGTGTGTTGCC 2160  
2199 AATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCGAGATGAGTGTGCTCGGAGA 2258  
2161 AATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCGAGATGAGTGTGCTCGGAGA 2220  
2259 GATACCGAGGCGAGTCCCAAGTCCCTTTAGGAGCAGGTCAACCCCTCTCTGACCTC 2318  
2221 GATAGCGAGGCGAGTCCCAAGTCCCTTTAGGAGCAGGTCAACCCCTCTCTGACCTC 2280  
2319 CAGCCATACATGGCCAGTTCCTTAAGCATCTGAGGATTCAGATGCACTGAGG 2378  
2281 CAGCCATACATGGCCAGTTCCTTAAGCATCTGAGGATTCAGATGCACTGAGG 2340  
2379 AACTCGGTGTATCGAGAGAGCATCTATGATGAGAGCAGCAGCGCTTTTGAC 2438  
2341 AACTCGGTGTATCGAGAGAGCATCTATGATGAGAGCAGCAGCGCTTTTGAC 2400  
2439 TTCTTCCTGCACTTCTGCTGCTCAGAGTGTGTAAGATTTGTGACAGTGTCTATGACAG 2498  
2401 TTCTTCCTGCACTTCTGCTGCTCAGAGTGTGTAAGATTTGTGACAGTGTCTATGACAG 2460  
2499 TGCCAGGGCATCCCGAGGCTCCAGCTATCCACCTGCTCTGAGTCTGTGTTGGA 2558  
2461 TGCCAGGGCATCCCGAGGCTCCAGCTATCCACCTGCTCTGAGTCTGTGTTGGA 2520  
2559 GACATCGAGAAACAGCTGTTTGTCTGAGGTGACAGCGGATGGGTTGCTTTACGTTTGT 2618

2521 GACATCGAGAAACAGCTGTTTGTCTGAGGTGACGGGATGGGTTGCTTTACGTTTGT 2580  
2619 GATGACTTTTCTGTTGGTGAAGCTTCACTTTGAGCAAGCAAAACCTTCTCTCAGACCTG 2678  
2581 GATGACTTTTCTGTTGGTGAAGCTTCACTTTGAGCAAGCAAAACCTTCTCTCAGACCTG 2640  
2679 GTCCATGGGTTCTCTGAGTATGGGTGATATAAATTTGACAGAGACAGTGTGTAACCTTC 2738  
2641 GTCCATGGGTTCTCTGAGTATGGGTGATATAAATTTGACAGAGACAGTGTGTAACCTTC 2700  
2739 CTTGTGGAGCCCTGGTACCTTGGGTGGTGCAGCTCCATACAGCTGCCCTGCTCACTGCTG 2798  
2701 CTTGTGGAGCCCTGGTACCTTGGGTGGTGCAGCTCCATACAGCTGCCCTGCTCACTGCTG 2760  
2799 TTTCCCTGCTGTGGCTTCTGCTGAGCACTCAGACTTTGGAGGTTCTGTGACTACTCA 2858  
2761 TTTCCCTGCTGTGGCTTCTGCTGAGCACTCAGACTTTGGAGGTTCTGTGACTACTCA 2820  
2859 GGTATGCCAGACCTCAATTAAAGACGAGCTTCACTTTCCAGAGTGTCTTCAAGCTGGG 2918  
2821 GGTATGCCAGACCTCAATTAAAGACGAGCTTCACTTTCCAGAGTGTCTTCAAGCTGGG 2880  
2919 AAGACCATGCGGAAACAGCTCCTGTGCTGCTTGGCTTGAAGTGTCAAGGTTCTATTTCTA 2978  
2881 AAGACCATGCGGAAACAGCTCCTGTGCTGCTTGGCTTGAAGTGTCAAGGTTCTATTTCTA 2940  
2979 GACTTGCAGTGAACAGCTCCAGACAGTCTCATCAATATATACAAGATCTTCTCTGCTT 3038  
2941 GACTTGCAGTGAACAGCTCCAGACAGTCTCATCAATATATACAAGATCTTCTCTGCTT 3000  
3039 CAGGCTTACAGGTTCCATGCTGTGATTCAGCTTCCCTTTTGACAGCGTGTAGGAG 3098  
3001 CAGGCTTACAGGTTCCATGCTGTGATTCAGCTTCCCTTTTGACAGCGTGTAGGAG 3060  
3099 AACCTCACATTTCTTCTGGGATCATCTCCAGCAAGCATCTGCTGTATCTATCTCTG 3158  
3061 AACCTCACATTTCTTCTGGGATCATCTCCAGCAAGCATCTGCTGTATCTATCTCTG 3120  
3159 AAGCTCAAGAAATCCAGGAATGACACTAAAGGCTCTGCTCTCTCTCTCTGAAAGCGCA 3218  
3121 AAGCTCAAGAAATCCAGGAATGACACTAAAGGCTCTGCTCTCTCTCTCTGAAAGCGCA 3180  
3219 CATTTGGCTTCTGCTACAGGCTTCTGCTCAAGCTGGCTGCTCATTTCTGTCACTACAAA 3278  
3181 CATTTGGCTTCTGCTACAGGCTTCTGCTCAAGCTGGCTGCTCATTTCTGTCACTACAAA 3240  
3279 TGTCTCTCTGGAGCTCTGAGGACAGCCCAAAACCTGCTGTGCGGAAGCTCCAGAGCG 3338  
3241 TGTCTCTCTGGAGCTCTGAGGACAGCCCAAAACCTGCTGTGCGGAAGCTCCAGAGCG 3300  
3339 ACAATGACCATCTTAAAGCTGAGCTGACCCAGCCCTAAGCAGACAGCTTTCAGACCAT 3398  
3301 ACAATGACCATCTTAAAGCTGAGCTGACCCAGCCCTAAGCAGACAGCTTTCAGACCAT 3360  
3399 TTGGACTAA 3407  
3361 TTGGACTAA 3369

RESULT 3  
AF149012  
LOCUS  
DEFINITION Mesocricetus auratus telomerase catalytic subunit mRNA, complete  
4170 bp mRNA linear ROD 21-FEB-2001  
ACCESSION AF149012  
VERSION AF149012.1  
KEYWORDS GI:6572639  
SOURCE Mesocricetus auratus [golden Syrian hamster]  
ORGANISM Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.

REFERENCE 1 (bases 1 to 4170)  
 AUTHORS Guo, W., Okamoto, M., Lee, Y.M., Baluda, M.A. and Park, N.H.  
 TITLE Enhanced activity of cloned hamster TERT gene promoter in transformed cells  
 JOURNAL Biochim. Biophys. Acta 1517 (3), 398-409 (2001)  
 MEDLINE 21240330  
 PUBMED 11342218  
 REFERENCE 2 (bases 1 to 4170)  
 AUTHORS Guo, W., Okamoto, M. and Park, N.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAY-1999) Dentistry, University of California, Los Angeles, 10833 Le Conte Ave., Los Angeles, CA 90095, USA  
 FEATURES Location/Qualifiers  
 source 1..4170

/organism="Mesocricetus auratus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10036"  
 /notes="common: Golden Syrian hamster"  
 28..3414  
 /notes="hamTERT"  
 /codon\_start=1  
 /product="telomerase catalytic subunit"  
 /protein\_id="AA17334.1"  
 /db\_xref="GI:567240"  
 /translation="MPAPRCNVAALLRSQYRQVWELATFVRLGPBGRQLVQEDP  
 KVFRLVACLVCVPWDSQPPADLSHOVSSKELVARVQRLBERGNVLTGFA  
 LVNAGGPPMTFTTSVRSYLPNSVTSRVSAGMILLNRGDDLLVYLARCALYL  
 LPPSCAYQVCSPLQYCAETWPSVSYRTPFVGRNTHLGSFVRNSSHQ  
 ANKPPPLPREAKRSLSTNRSVPSKARCDLAPLEKGPYQAVTPSPDXTWPNP  
 AKSHAVPISRTTKEDLSGKAPGLSRSGSVCKPSPSTLSQSLQNAFQRLPYTE  
 TKRFLVSRGGQLNPNFLNNLPSTLGGARRLVEILPLGMRPTSGPLCGRRLSK  
 RYWMQLPQQLLVNARCPVRLLRSHRCPRTAAHQVAGALNTTSPORLNLRLHS  
 SPQVYVFLQACVGLQVPLVGLSGHNRQREFKVAKEFISLCKYKXLSQELTWKXY  
 QDCWRLSRSPGNVCPAAEHTRERILAFELFLMDAYVVELLSFTYETTFQNR  
 LFFYNSMWRRLQSLGVRRHLRERLQELSOEVRQOEAWPAMPICRURFIPKESGL  
 RTFVNVNLTADPDKQAOHFTQCLTFLSVNLKELTHTNLGASVLGNDLYRT  
 WPTFLVRLTPRPFVADVTGAYDAIPQKLVEVTIANMRHPDMSYCIHQYAV  
 VDRDQQLHKSFQVNSTLSDLPQHMGEFLKLDSDTSALRNSVVIQSLNSLEAS  
 SSLFDPLFVANSVVKGGYVQCGIPOGSLSTLLCLSCFDMENKLPFAEVOO  
 LLRLREDDFLVTLPLVQABAFRALVRGIPYEGCMINIKQTVNFPVDAGTLQGT  
 PHLLPACILFPACGGLLLOTLEVLCDYTVARTSIKASLTFTQPKAGNNRQKLLA  
 VLRLKCHSLFLDLQNSLQTVINVKIPLLOAYRFACALQLPFDQHVRLNPAFLS  
 LISNIAASCSTILKVNAGMTLKAAGSGFPFPAARWUCITQAFLLKAGHSVITKCL  
 LGPLRTAQKQCRKLPAPRAWTMAILEATADPALSTDQFILD"

ASE COUNT 927 a 1177 c 1126 g 940 t  
 RIGIN

Query Match 73.8%; Score 2578.6; DB 10; Length 4170;  
 Best Local Similarity 85.4%; Pred. No. 0;  
 Matches 2941, Conservative 0; Mismatches 479; Indels 25; Gaps 5;  
 12 GGGAGGCCATCCCGCCCTTGACACATGACCCCGCTCTCTGTCGCCCGCGTGGCG 71  
 1 GGGAGGCCCGCGCGGATCTTGACCGGATGCGCCCGCGCGCCCTTCCGCGCGTGGCG 60  
 72 TCTCTGTCGCGACCGATACCGGAGGTGTGGCGCTGGCAACTTGTGTGGCGCGCTG 131  
 61 GCTCTGTCGCGAGTCAATACCGTCAAGTGTGGCGCTGGCAACTTGTGTGGCGCGCTG 120  
 132 GGGCCGAGGCGAGCGGCTGTGCAACCCGGGACCGGAGATCAACGCACTTGTGT 191  
 121 GGAATGAGGCGAGGAGCTGTGCAACCCGGGACCGGAGATCAACGCACTTGTGT 180  
 192 GCCCAATGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 251  
 181 GCCCGGTGCTAGT 240  
 252 CACGAGGTGTCT 311  
 241 CACGAGGTGTCT 300  
 312 AACGAGAGAAAGT 371

301 GCGAGAGAGAAAGT 360  
 372 CCCATGGCTTCACTAGT 431  
 361 CCATGACATTCACACACGCGTGGCAGTACTCTGCCAATCTCGGTGACTGTCTCTG 420  
 432 COTGTCAGT 491  
 421 CGCGTCAGT 480  
 492 CTGCTGGGACACTGT 551  
 481 CTGCTGGGCGCTGT 540  
 552 GGGTCTCCCTGTACCAAAATTTGTGCGCAGATATCTGGCCCTCTGTGTGTGTGTGT 611  
 541 GGTTCACCCCTGTACCAAAATTTGTGCGCAGATATCTGGCCCTCTGTGTGTGTGTGT 600  
 612 TACAGGCCCAACCCGACCGGTGGGAGAAATTTTCACTAAGCTTAGTGTCTTACACAGATC 671  
 601 TACAGGCCCAACAGACCGGTGGGAGAAATTTTCTACTCATCTTGGATCCACACACCGGTC 660  
 672 AAGACGACTGT 731  
 661 AGAAGACGACTGT 720  
 732 AGGCATCTGAGTCTCAACAGTACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791  
 721 CGGAGTCTAAGCATCAACATAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
 792 GTCCCGAGAGT 851  
 781 GCGCGGAGACTGT 840  
 852 TGGGTGCGCAAGTCTGT 902  
 841 TGGGTGCGCAAAATCTGT 900  
 903 TTGTCTTCTTAAAGAAAGT 962  
 901 TTGTCTTCCGGGT 960  
 963 AAGCCAGCTTCAATCTGT 1022  
 961 AAGCCAGCTTCAATCTGT 1020  
 1023 TTTATTGAGACCAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1082  
 1021 TATATCTGAGACCAAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
 1083 TCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1142  
 1081 TGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140  
 1143 ATCTTTTGGGCTAAGGCTTAGACATCAGGACCACTGTGTGTGTGTGTGTGTGTGTGT 1202  
 1141 CTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
 1203 CGTGCATCTGT 1262  
 1201 AAGCGCTACTGT 1260  
 1263 CAATATGTCTGACTCTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1322  
 1261 CGGTATGTCTGCTCTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320  
 1323 GATGCTGTGAA---CACCAGGCCACCGCCTCATGGAATTTGTGTGTGTGTGTGTGTGT 1379  
 1321 GGTGCTTGAACACACACCGCCTCATGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380  
 1380 CCTTGGCAGGTATATGTTTCTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1439  
 1381 CCTTGGCAGGTATATGTTTCTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440



```
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1758)
Mong,S., Gao,S., Xu,X. and Yu,H.
Rat telomerase catalytic subunit, rTERT
Unpublished
2 (bases 1 to 1758)
Wong,S., Gao,S., Xu,X. and Yu,H.
Direct Submission
Submitted (21-MAR-2000) IMRB, National University of Singapore, 10
Medical Drive 117597, Singapore
Location/Qualifiers
1..1758
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
<1..1728
/notes="rTERT"
/codon_start=1
/product="telomerase catalytic subunit"
/protein_id="AAF62177.1"
/db_xref="GI:7407653"
/translation="RSFFVITESTQKURLFFYRKSVMKLSIGVROHLERVELREL
SDEVRHHQDTWLMPIELRLPIKNGRLPIVMSYSMTGRALGRKQAHPTORLK
TIFSMUNYERTKHPHLMSSVIGMDIYRTWRAFLVRALDQTPMYFYKADVTGY
DAIPOKLVENVNMRHSESTYICRQAVVRDSQGVHKSFRQVITLSDLPNG
QFLKHLQDSALRNSVIEQSIEMNESSLSDFLFLHRSVIVKQRCYTCQG
IPOGSLHLSLSCFGDMENKLFVAVRDGLLRFVDDFLVFLPHLDQKATFLSTV
HGVEYGMINKLQTVNPFVFGTGLGAAPQYPAHCLFPWCGLLDQTLVEFCDY
SGVAQNSIKTSITQSVFKAGTKMKNLLSVLRKCHGLFDLDQVNSLQTVCNINYKI
FLQAVRFACVIOIOPFDQVKNLTFPLGIHISQASCYAILKVNPGMTLKASGSF
PPEAAHWLCTQFLKLAHSAVYIKLLGLPLNTAQKLLCRKLPEATWILKAAADPAL
STDQFQILDD"
misc_feature
<1..87
/notes="Region: motif T"
SE COUNT 398 a 486 c 447 g 427 t
IGIN
Query Match 50.2%; Score 1756; DB 10; Length 1758;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1680 AGGTCATCTTTTACATCAGAGAGACATTCAGAGAACAGGCTCTTCTTACCGT 1739
1 AGGTCATCTTTTACATCAGAGAGACATTCAGAGAACAGGCTCTTCTTACCGT 60
1740 AAGAGTGTGGAGCAAGCTGCGAGAGCATGGAGTCAGGCAACATTCAGAGAGTGGG 1739
61 AAGAGTGTGGAGCAAGCTGCGAGAGCATGGAGTCAGGCAACATTCAGAGAGTGGG 120
1800 CTACGGAGCTGTCAAGAGGAGGTTCAGGATCAGGACACCTGGGTAGCCATGCC 1859
121 CTACGGAGCTGTCAAGAGGAGGTTCAGGATCAGGACACCTGGGTAGCCATGCC 180
1860 ATCTGCAGATGGCGTTTATCCCAAGCCCAACGGCTGGCCCAATTTGTGAACATGAT 1919
181 ATCTGCAGATGGCGTTTATCCCAAGCCCAACGGCTGGCCCAATTTGTGAACATGAT 240
1920 TATAGCATGGGTACCAAGAGCTTTGGGCAAGAGCAAGCAGGCCCATTTACCCAGAGT 1979
241 TATAGCATGGGTACCAAGAGCTTTGGGCAAGAGCAAGCAGGCCCATTTACCCAGAGT 300
1980 CTCAGAGCTCTCTTTCAGCATGCTCAATATGAGCGGACAAAACATCTCCTATGGGG 2039
301 CTCAGAGCTCTCTTTCAGCATGCTCAATATGAGCGGACAAAACATCTCCTATGGGG 360
2040 TCTTCTGTACTGGGTATGATGATCTACAGACCTGGCGGCCCTTTGTGTCGGTGTG 2099
361 TCTTCTGTACTGGGTATGATGATCTACAGACCTGGCGGCCCTTTGTGTCGGTGTG 420
2100 CGTGTCTGGACACAGACCCAGGATGTATTTGTTAAGCAGATGTGACCGGGCCTAT 2159
```



/transl_table=11	
/product="telomerase catalytic subunit"	
/protein_id="AAC40212.1"	
/db_xref="GI:3023055"	
/translation="MPRAPRCRAVSLRSLRSHYREVLPATFVRRLVRLGQGWRLVQRGDP	
AATFRLVAOCLVCPMDAPFPAPPAQSPQVCLKELVARLVORLCEKAGKVLAAFGFA	
LDGASGGPEAFTTVSRVSLVNTVTDALRGSGAGLLLRVDDVLVHLLARCLPA	
LVAPSCAYQVCPPLYLQAAATCAAPPHPASGPRRLRCERANHSVREAGVPLGLPA	
PGARRGGASRLPLFPRGRGAPEPRTVPQSGWAHFGRTGSDRGFCVWSPFA	
RPAEATSLGEGALSGFRSHSPVQGHAGPPSPRPMDTFCPPVATATKFLHYS	
SGDKQLRPSFLLSRLPTLGARLVETIFLGRPMPEGTFRLLPRLPQRYWQMRPL	
FLELLGNHACQPGVLLKTHCLTAAVTPAAGVCAREKPGQSAAPSEEDTDPRLVQ	
LHLQHSQPMQVTFVRACLRVLPGLSGRHNERRFLRNKKFISLGHAKLSLOSL	
TWKSVGCAKMLRRSPGCVGPAASHRLEETLAKFLHMLNSVVVVELLASFVYVTE	
TFQKNRLFFVRKSVMSKLSQSIGIRGHLXRVQLRELSEAVRQHREANPALLTSRLFI	
PKDDLRWPTVNMVDYVGAATPPAREKGRSLRVRKALFVUNYERAKRPGLLGASVLG	
LDDIHRAMPTFVLVRVAQDPPELVFVVDVTGAYDTIPQRLTEVASIKPQNTVC	
VRAYVYQKAAGHVGAFAKFSVSTLTDLPYMRQFVAHLQETSPTRDADVIEOSSLS	
NBSAGSLDFVFLRFGHCHAVIRLGSYVQCQIQPGQSLSTLLGSLCVGMENKLFAG	
IRDGLLILAVDDLFLVPLHTLTAFLRTLVIRGVEYGVVNLKTKVWNPVFDEAL	
GTAFAVMPAHGLFPWCGLLDDTRILEVQSDYSVARTSIRASLTENRGRKAGNMRRI	
KLGVRLKCHSLFLDLVNSLQTVCTWNTKYLILQAVRFAVLQLPHQVQWKNPT	
FFLRLSDTRASLCYSITLAKNAGMSLGAKGAAPLFPSEAVOWLCHQAFLLKLTFRHVT	
YVPLLSLRTPAQTLRSKLPJLITALEAANPALPUSDFKTLIDPYDVPDIAGPYD	
VPDYAG"	
BASE COUNT	1886 a 2551 c 2484 g 2038 t 1 others
ORIGIN	
Query Match	45.9%; Score 1603.6; DB 12; Length 8960;
Best Local Similarity	68.7%; Pred. No. 0;
Matches 2389; Conservative	0; Mismatches 979; Indels 108; Gaps 9;
QY	23 CCGCGCCTTGAGCAATACCCCGCCTCTCTGTTGCCCGCGTGCCTCTCTGCTCGC 82
DB	1131 CCGGECACCCCGGATCGCGCTCCCGCTGCCGAGCGTCCCTGCTCGC 1190
QY	83 CAGCGGATACCGGAGGTGTGGCGCTGTGGCACTTTGTGTGGCGGCTGTGGGCGCCGAGGG 142
DB	1191 CAGCCACTACCGCAGGTGTGCGCTGTGCACCTGTTCGTGTGGCGGCTGTGGGCGCCGAGGG 1250
QY	143 CAGCGGGCTTGTGCACCCCGGGACCCGAGATCTACCGCACTTTGTGTGCCCAATGCT 202
DB	1251 CTGGCGCTGTGTGACGCGCGGACCCGCGGCTTCCGCGGCTGTGTGGCCAGTGCCT 1310
QY	203 AGTGTGCATGCATGGGGCTCACAGCCTCAACTGCGGACCTTCTCTTCCACCAAGTGT 262
DB	1311 GGTGTGTGCGCTGTGGACACAGCGCGGCGCCCGCGCGCCCTCTTCCGCGAGGTGT 1370
QY	263 ATCCCTGAAGAGCTGTGTGCGCGAGGTGTGTGAGAGACTGTGCGAGCGCACAGAGAGAA 322
DB	1371 GTGCTGAAGAGGTGTGTGCGCGAGTGTGTGAGAGGCTGTGCGAGCGCGCGCGGAGAA 1430
QY	323 CGTGTGCTCTTTTGGCTTTTGAGCTGTCTTAACGAGGCGCAGAGCGGGCTTCCCATGGCCTT 382
DB	1431 CGTGTGCGCTTTCGCTTTCGCTGTGTGAAGGGGCGCGCGGGGCGCCCGCGAGGCGCTT 1490
QY	383 CACTGTGTAGCTGTGCTGTACTTGTGCCAACACTGTATTGAGACCTTCGCTGTGTCACTGG 442
DB	1491 CACCACCAAGCTGTGCGAGCTACCTGTGCCAACACGGTACCGGACGCACTGTGCGGGGAGCGG 1550
QY	443 TGCATGGATGCTACTGTGTGAGCGAGTGTGGCGACGACCTGTCTGTCTACTCTGTGGCACA 502
DB	1551 GCGTGTGGGGCTGTGTGTGCGCGGTGTGGGACGACGTGTGTTTCACTTGTGTGGCAGC 1610
QY	503 CTGTGCTCTTTATCTTCTGTGTGCCCGCAGCTGTGCTACCAAGTGTGTGGGTCTCCCT 562
DB	1611 CTGCGCGCTCTTTTGTGTGTGTGCTCCAGCTGCGCTTACCAAGTGTGTGGGCGCCGCT 1670
QY	563 GTACCAAAATTTGTGCCACCAACGATATCTGGCCCTCTGTGTCCGCTAGTTACAGGCCCAAC 622
DB	1671 GTACCAAGCTGTGCGCTGTCACTAGGCCCGGCGCCCGCACAGCTGTGTGACACCCGAG 1730
QY	623 CCGACCCGTGGCGAGGAATTTCACTAACTTGTAGTTCTTTCACACAGATCAAGACAGTAG 682



1731 GCGTCTGGGATGC-----GAAAGCGGCTGGAACCATAG 1763  
683 TCGCAGAGACACGAAACCCCTGGGCTTGGCATCTCGAGGTACAAAGAGGCATCTGAG 742  
1764 CGTCAGGAGGCGCGGGTCCCTTGGGCTTGGCAGCCCGGGTGGAGAGGCGCGGGGG 1823  
743 TCTCAGCAGTCAAGTGTGCTTCAAGTAAAGAGGCGAGATCTATCTCTGCTCCGAGAGT 802  
1824 CAGTGCACAGCCGAAGTCTGCGGTGTGCGCAAGAGGCGCGAGCGTGGCGTCCCTTGAGCC 1883  
803 GAGAGAGGACCCACAGCAGGTGTCTACCAACCCATCAGGCAATCATGSGTGGCAAG 862  
1884 GAGCGGACGCGCGTGGCGAGGGGTCTCGGCGCCACCGGCGAGGACGCGTGGACGAG 1943  
863 TCTGCTGGTCCCGGAGGTGCTCTACT-----GCAGAGAAAGATTGTTCTTTAA 913  
1944 TGACCGTGGTCTTGTGTGGTGTCACTGCGCAGACCCCGCGAAGAGCCACTCTTTGGA 2003  
914 AGGAAAGTGTCTGACCTGAGTCTCTC---TGSGTGGTGTGCTGTAAACACAAAGCCAG 970  
2004 GSGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGGCGCCAGCACCGCGGG 2063  
971 CTCCTCAT---CTGCTGTACACACCCCGCAAAATGCTTTACGCTCAGCCATTAT 1027  
2064 CCCCCATCCATCGCGGCCACACGCTCTCTGGACACGCTTGTCTCCCGGTGTAGC 2123  
1028 TGAGACGAGATTTCTTTACTCCAGGGGAGATGCCAAGAGGCTCTAAACCCCTCAT 1087  
2124 CGAGACCAAGCACTCTCTCTACTCTCTCAGCGCA---CAAGAGGACAGCTGCGGCTCTCT 2180  
1088 CTTACTCAGCAACCTCCAGCTTAATCTGCTGGGCGAGGAGACTGGTGGAGATCATCTT 1147  
2181 CTTACTCAGCTCTCTGAGGCCCGAGCTCTGCTGGAGGCTCTGTGAGACCATCTT 2240  
1148 TCTGGGCTCAGGCTTAGCAGATCAGAGCACCTCTGACGACACACCTCTATGGGTG 1207  
2241 TCTGGGTTCCAGGCGCTGTATGCGCAGGAGCTCCCGCAGGTTGCGCCCGCTGCGCCAGCG 2300  
1208 ATACTGGCAGATGGCGGCTCTTCTCAAGAGCTGCTGTGTGAACCAATGAGAGTGGCAATA 1267  
2301 CTACTGGCAATGGCGGCTCTTCTGAGGCTGCTTGGGAACCAAGCGGAGTGGCCCTA 2360  
1268 TGTGAGCTCTCAGGTCACATTCAGGTTTCCAAAGCAGCAACCAAGAGTGCACAGATGC 1327  
2361 CGGGTGTCTCTAAGACGCACTCCCGCTGCGAGCTGCGGTGACCCAGCAGCGGCTGT 2420  
1328 CT-----TGAAACCAAGCC 1342  
2421 CTGTGCGCGGAGAGCCCGCAGGCTCTGTGGGCGCCCGCGAGGAGGAGACAGAGCC 2480  
1343 ACCGACCTCATGATTTGCTCCGCTGACAGAGTCTCTGCGAGGTATATGGTTTCT 1402  
2481 CGTCCGCTGTGTGAGCTGCTCCGCGACACAGAGCCCTCTGCGAGGTGTACGCTTCGT 2540  
1403 TCGGCGCTGTCTCGAAGTGTGTCTGTAGTCTCTGGGTACAGGCAATGAGCG 1462  
2541 GCGGCGCTGTCTGCGCGCTGTGTGCGGCTCTGCGGCTCTCGAGCAACAGBAAC 2600  
1463 CGGCTTTCTTAAGAACTTAAAGAAAGTTCACTCTCGTTGGGGAATAACGGCAAGCTATCACT 1522  
2601 CGGCTTCTCAGAAACCAAGAAAGTTCACTCTCGTTGGGGAAGCATGCAAGCTCTCGCT 2660  
1523 GCAGGAGCTGTGCGAGATGAAAGTGAAGATTCGCTCGCTCCGCGAGCGCCGGG 1582  
2661 GCAGGAGCTGAGCTGGAAGATGAGCTGTGGGCTGCGCTTGGCTGCGGAGGCGCCAGG 2720  
1583 GAAGGACCGTGTCCCGCTGAGAGGACCGCTGTGAGGAGAGATCTCTGGGTAGTTCCT 1642  
2721 GGTGCTGTGTTCGCGCGCGAGAGACCGCTCTGCTGAGGAGATCTCTGGGCAAGTTCT 2780  
1643 GTTCTGCTGTGAGACATACGTGTGTGAGCTGTAGTCTATCTTTTACATCAGAGA 1702

DG 2781 GCACTGGCTGATGATGTGTAGTGTGTCAGGTCTCTCAGGTCTTTCTTTTATGTACACGA 2840  
QY 1703 GAGCAGATTCCAGAGAGACAGGCTCTTCTCTACCGTAAAGTGTGTGAGCAAGCTGCA 1762  
DB 2841 GACCACTTTCAAAGAAACAGGCTCTTTCTACCGAAGAGTGTCTGGAGCAATTTGCA 2900  
QY 1763 GAGCATTGAGGTGAGGCAACACTTTGAGAGAGTGGGCTACGGAGCTGTCTCAAGAGGA 1822  
DB 2901 AAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTGGAAGCAGA 2960  
QY 1823 GGTGAGCATCACCAGGACACCTGCTAGCCATCTGAGACTGCGCTTCTCATCCC 1882  
DB 2961 GGTGAGGAGCATCGGAGAGCGAGCCCGCTCTGAGCTGCCAGACTCGCTTCTATCCC 3020  
QY 1883 CAAGCCCAAGCGCTCGGCGCCATGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
DB 3021 CAAGCTGAGCGGCTGCGGCGGATTTGGAACATGGAATAGCTGCTGGGAGCCAGAACGTT 3080  
QY 1943 GGGCAGAGAAAGCAGGCGCCAGCATTTCAACCAAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
DB 3081 CCGCAGAGAAAGAGCGCGGAGCGTCTCACTCGAGGGTGAAGGCACTGTTCAAGCGTGT 3140  
QY 2003 CAACTATGAGCGGACAAACATCTCTACCTATGGGCTCTCTGACTGGGTATGAATGA 2062  
DB 3141 CAATGAGAGCGGCGCGGCGCTCTGCGGCTCTGCGGCGCTCTGCTGGGCTGAGAGA 3200  
QY 2063 CATCTACAGGACTCGCGGCGCTTGTGTGCTGGTGTGCTGTGAGACCAACACCCAG 2122  
DB 3201 TATCCACAGGCGCTGCGGCGCTCTGCTGCTGCTGCTGCGGCGCCAGGACCCCGGCTCTGA 3260  
QY 2123 GATGTACTTTGTTAAGGAGAGATGTGACCGGCGGCTATGATGTCATCCCCCGAGGTAAAGCT 2182  
DB 3261 GCTGTACTTTGTCAAGTGGATGTGACGCGGCGGTACGACACCACTCCCGCAGCAGGCT 3320  
QY 2183 GGTGAGGTTGTGCAATATGATCAGGCACTCGAGAGCACTGACTGTATCGGCGAGTA 2242  
DB 3321 CAGGAGGTTCTATCGCGCGATCTATCAAAACCC---AGAACACGTTCTGCTGCGTGGTA 3377  
QY 2243 TGCAGTGCTCGGAGAGATAGCCAAAGCCCAAGTCCAAAGTCTTTTAGGAGACAGGTCAAC 2302  
DB 3378 TGCGGTGTCTCAGAAAGCGCGCCATAGGCGCACTGCGCAAGGCGCTTCAAGAGCCAGTCTC 3437  
QY 2303 CACCTCTCTGAGCTCCAGCCATACATGAGGCGAGTCTCTTAAGCATCTGACAGATTCAGA 2362  
DB 3438 TACTTGAACAGACCTTCAGCGGTATCATGCGAGTGTGCTGCTCAGCTGAG---GA 3491  
QY 2363 TGCAGTGCTGAGGAACTCGGTTGTCTGAGTGTGAGAGAGCATCTCTATGAAATGAGAGAG 2422  
DB 3492 GACCAAGCGCTGAGGAGATGCGGCTGCTCATGAGAGAGCTCTCTCTGAAATGAGGCGAG 3551  
QY 2423 CAGCAGCTGTTTGAATCTTCTCTGCACTCTCTGCGCTCAAGTGTCTGAAAGATTTGTTGA 2482  
DB 3552 CAGTGGCTCTTCTGAGCTCTTCTTACGCTTCTTACGCTTCTGCGAGCGGCTGCGCATCAGGG 3611  
QY 2483 CAGGTCTATACGAGTGGCGAGGCGATCCCGAGGCTCCAGGCTATCCAGCTGCTCTG 2542  
DB 3612 CAGTCTTACGTTCTGAGTGGCGAGGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 3671  
QY 2543 CAGTCTGTTTTCGAGACATGAGAAACAAGCTGTTTGTGAGGTTGAGCGGATGAGGTT 2602  
DB 3672 CAGCTGTCTACGCGGAGATGAGAAACAAGCTGTTTGGGGGATTCGCGGAGCGGCT 3731  
QY 2603 GCTTTTACGTTTGTGATGATCTTCTGTTGGTGGAGCTCTCACTTGGAGCAAGCAAAAC 2662  
DB 3732 GCTCTGCGTTTGTGAGTATGATTTCTTGTGTTGAGACCTCTCACTCAGCCAGCGCAAAAC 3791  
QY 2663 CTTCTCTACGACCTCTGTTTCCATGCGGTTCTGAGTATGAGTGTGATTAACCTTGCAGAA 2722  
DB 3792 CTTCTCTCAGGACCTCTGTTCCGAGGTTCTCTGAGTATGAGTGTGAGTGTGAGTGTGAG 3851  
QY 2723 GACAGTGTGAACTTCTCTCTGTTGGAGCTGTGTAACCTGGGTGTGAGCTTCCATACAGCT 2782  
DB 3852 GACAGTGTGAACTTCTCTCTGTTAGAAAGCAGAGGCGCTTGGGTGGCAGCGCTTTTGTTCAGAT 3911



1148 TCTGGGCTCAAGCCCTAGAGCACTCAGAGCACTCTGCGAGGACACACCGTCTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCCCTGGATGACGAGGACTCCC GCGAGGTGCGCCGCTGCGCCAGCG 1209  
1208 ATATGCGCAGATCGGCCCTGTGTTCCAAACAGTGTGTTGTAACCATGACAGTGCCAATA 1267  
1210 CTAATGGCAATCGGCCCTGTGTTCTGAGAGCTGTCTGGGAACACAGCGAGTGCCTCTA 1269  
1268 TGT CAGACTCTCAGGTTCATGTCAGGTTTCGAACAGCAAAACAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAGACGCACTGCGCGCTGCGAGCTGCGGTACCCACGACGCCGTGT 1329  
1328 CT-----TGAACACACAGCCC 1342  
1330 CTGTGCGCGGAGAGCCCGCAGGGCTCTGTGCGGCCCGCCGAGAGGAGGACACAGACCC 1389  
1343 ACCGCACTCATGATATTGCTCCGCTGACAGCAGTCCCTGGCAGGTATATGGTTTTCT 1402  
1390 CGCTGCTGTGTCAGCTGCTCCGCGCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGT 1449  
1403 TCGGGCTGTCTCTGCAAGGTGTGCTGCTAGTCTCTGGGGTACACAGGACCAATGAGCG 1462  
1450 CGGGCTGCTCTGCGCGGTGTGCCCCAGGCTCTGCGGGCTCCAGGCACAAAGACG 1509  
1463 CGCTCTTTAGAACTTTAAGAGTTCACTCTGTTGGGAAATACGGCAAGTATCATCT 1522  
1510 CGCTTCTCTAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTTCGT 1569  
1523 GCAGAACTGATGTGGGAAGATGAAGTAGAGATTGCCACTGGCTCGCGACAGCCCGGG 1582  
1570 GCAGAGCTGACGTGGAAGATGAGCGTGGGACCTGCGTTGGCTGGCAGAGCCCGAG 1629  
1583 GAAGACCGTGTCCCGCTCCAGAGCACCGTCTGAGGGAGAGGATCCGTGGCTACGTTCCCT 1642  
1630 GGTGTGCTGTGTTCCGCGCGCAGAGCACCGTCTGCTGAGGAGATCTCGGCAAGTTCCCT 1689  
1643 GTTCTGGCTGATGAGCACATACCTGTGTGACGTGTTAGGTCACTCTTTTACATCACAGA 1702  
1690 GCATCTGGCTGATGATGTGTACCTCTCGCAGCTGCTCAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCACATTCAGAAAGAACAGGCTCTTCTTCTACCGTAAAGAGTGTGGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGGAGTCAGCAACAACCTTGAGAGAGTGTGGGCTAGCGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGAATACAGACAGCACTTGAAGAGGTGTGAGCTGTGCGGAGCTGTCTGGAGCAGA 1869  
1823 GTTCAGGCATCACAGGACACCTGGCTAGCCATGCCCCATCTGCAAGACTGCGCTTCATCCC 1882  
1870 GGTACGGCAGCATCGGGNAGCCAGGCCCGCCCTGTGTACGTCCAGACTCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTGCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAGGCTGACGGCTGTGGCGGATTTGTGAACATGAGCTACGTCGTGTGGAGCCAGAAAGTT 1989  
1943 GGGCAGAGAGGACAGGCCCAGCAATTCACACCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CGCGCAGAAAGAGGCCCGAGCGTCTACCTCGAGGGTGAAGGCACTGATTCAGCGTGT 2049  
2003 CAACCTATAGCGGCAAAAAATCCTCACCTTATGGGGTCTCTGTACTGGGTATGAATGA 2062  
2050 CAACATCAGCGCGCGGGCGCCCGCCCTCTGTGGCGGCTCTGTGTCTGGCCTGTGACGA 2109  
2063 CATCTACAGGACCTGGCGGCTTTGTGTGCGGTGTGCGTGTCTGTGAGCAGACACCCAG 2122  
2110 TATCCACAGGCTGTGGCGACCTTCTGTGTGCGGTGTGCGGGCCAGACCCCGCCCTGA 2169  
2123 GATGATCTTTGTTAAGCAGATGTGACCGGGCGCTATGATGCGCATCCCCAGGGGTAACT 2182  
2170 GCTGTACTTTGTCAAGGTGTGATGACCGGGCGGTATCGACACCATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTCTGTCCAAATATGATCAGGCACTCGGAGAGCACTGATCTATTCGCCCACTA 2242

2230 CACGGAGGTCAATGCCAGACATCATCA--AACCCAGAAACAGCTACTCGGTGCGTCGGTTA 2286

2243 TGCAGTGGTCCGGAGAGATAGCGAAGGCCAAGTCCAAGTCCCTTTTAGBAGACAGGTCAC 2302

2287 TCGCGTGGTCCAGAAGCGCCCATGGGCACTCGCGAAGGCCTTCAAGAGCCACGTCCTC 2346

2303 CACCTCTCTGACCTCCAGGCCATATACATGGCCAGTTCCTTAAGCATCTGACGAGATTGAGA 2362

2347 TACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGCTGCTCACCTGCAG-----GA 2400

2363 TGCCAGTGGCATGAGGAATCCCGTTGTCTATCCAGCAGAGCATCTCTATGATGAGACGAG 2422

2401 GACCAGCCCGCTGAGGATGCGCGTCTCATCAGCAGAGCTCTCTCCCTGAATGAGCCAG 2460

2423 CAGCAGCCTGTTTGACATTTCTCTGCACTTCCTCGGTCAACAGTGTGCTAAAGATTGTTGA 2482

2461 CAGTGGCGCTTTGACGCTCTTCAAGCTTCTATGTCACCAACAGCCGCGTGCATCAGGGG 2520

2483 CAGGTGCTATACGACAGTGCACAGGGCATCCCCAGAGGCTCCAGCCTATCCACCTGCTCTG 2542

2521 CAGGTCTCACTCGTCCAGTGCACAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2580

2543 CAGTCTGTGTTTCGGAGACATGAGAAACAGCTGTTTCTGTGAGGTGCAGCGGATGCGGTT 2602

2581 CAGCCTGTGCTACCGCGACATGAGAAACAGCTGTTTTCGGGGATTCGGCGGAGCGGCT 2640

2603 GCTTTACGTTTGTGTGACATTTCTGTTGTGACGCTCACTTTGGACCAAGCAAAAC 2662

2641 GCTCTCGGTTTGGTGAGATGATTTCTGTTGTGTGACACTCTCACCTCAACCCAGCGGAAAC 2700

2663 CTTCCTCAGCACCTCGTCTCATGGCGCTTCTCAGTATGGGTGCATGATAAACTTTCAGAA 2722

2701 CTTCCTCAGGACCTCTGTCGAGGTGTCCCTCAGTATGCTGCGTGGTGAATTCGCGAA 2760

2723 GACAGTGGTGAACCTTCCTCTGAGAGCCTGGTACCTCGGTGGTGCAGCTCCATCAGCT 2782

2761 GACAGTGGTGAACCTTCCTCTGAGAGCAGGACCTGGGTGGCAGCGCTTTTGTTCAGAT 2820

2783 GCCTGCTCACTGCGCTTTCCCTGGTGTGGCTTGCTGTGGACACTCAGACTTTGGAGGT 2842

2821 GCGCGCCACCGCCCTATTCGCCCTGGTGGCGGCTGTGTGTGATACCCGACCCCTGGAGGT 2880

2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGAGAGCCCTCACCTCCAGAG 2902

2881 GCAGAGCGACTATCTCAGCTATGCCCGGACCTTCATCAGAGCCAGTCTCACCTTCAACCG 2940

2903 TGTCTTTAAGCTGGGAAACCATGCGAAACAGGCTCCTGTGCGTCTTGGGTTGAAGTG 2962

2941 CGGCTTCAGGCTGGGAGAACATGGCTCGCAAACTCTTTGGGGTCTTGGCGCTGAAGTG 3000

2963 TCACGGCTCTATTTCTAGACTTGCAAGTGAAACAGCCTCCAGACAGCTGTGCATCAATATA 3022

3001 TCACAGCCCTGTTTCTGCAATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA 3060

3023 CAAAGTCTTCCTCGGCTTCAGGCTTACAGGTTCCATGCATGTGTGATCAGCTTCCCTTTGA 3082

3061 CAGGATCTCTCTGCTCGAGCGTACAGGTTTCAAGCATGTGTGCTGAGCTCTCCCATTTCA 3120

3083 CCAGCGTGTTAGGAAGACCTCACTTTCTTGGGCATCATCTCCAGCCGAAGCATCTGT 3142

3121 TCGCAAGTTTGAAGAACCCCACTTTTCTCGCGCGTCACTCTGACACGGGCTCCCT 3180

3143 CTGCTATGCTATCCTGAAGTCAAGAAATCCAGGAATGACATTAAG-----GGCTC 3193

3181 CTGCTACTCCATCTGAAAGCCAGAAACGACGGGATGTGGCTGGGGGCAAGGGCGCGCG 3240

3194 TGGCTCCTTTCTCTGAAGCGGCACTTTGGCTCTGTACACAGGCTTCTCTGTCTCAAGCT 3253

3241 CGGCGCTCTGCGCTCCAGGCGCGTGAAGTGTGTGTGCCACCAAGCATTTCTGTCTAAGCT 3300

3254 GCGTCTCATTTCTCATCTACAAATGTCTCTCTGGGACCTCTGAGGACAGGCCCAAAACT 3313

3301 GACTCGACACCGTGTACCTACCTAGTGCACCTCTCTGGGTCACTCAGGACAGCCAGACGCA 3360  
3314 GCTGTGCGGGAAGTCCCGAGGAGGACATAGACATCTTTAAAGCTGAGAGTGAACCCAGC 3373  
3361 GCTGAGTCGGAAGTCCCGGAGCGAGCTGACTGCTGGAGCGCGACCAACCCGCG 3420  
3374 CTTAAGCACAGACTTTTCAGACCACTTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTTCAGACCACTCTCTGGACTGA 3454

UL7 7  
75848  
US ARI75848 4015 bp DNA linear PAT 17-DBC-2001  
INITION Sequence 224 from patent US 6309867.  
SSION ARI75848  
WORDS ARI75848.1 GI:17917147  
Unknown.  
RCE Unknown.  
RGANISM Unclassified.  
1 (bases 1 to 4015)  
BRENC Eech,T.R. and Nakamura,T.  
UTHORS Telomerase  
FILE Patent: US 6309867-A 224 30-OCT-2001;  
URNAL Location/Qualifiers  
TURES 1. 4015  
source /organism="unknown"  
E COUNT 663 a 1363 c 1275 g 714 t  
GIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGCCCTTGAGCAATAGACCCCGCTCTCTGTTGCCCGCGGTGCTCTCTGCTGG 82  
40 CCCGCCAACCCCGGATGCGCGCGCTCCCGCTGCCGAGCCGTGGCTCCCTGCTGG 99  
83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACTTTGTGCGCGCTGGGGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGCGCGCTGGCAACTTTGTGCGCGCTGGGGCCGAGGG 159  
143 CAGCGGCTTGTGCAACCGGAGACCGGAGATCTACCGCACTTTGTTGCGCAATGCT 202  
160 CTGGCGGTGTGTGACGCGCGGGACCGCGGCTTTCCGCGCTGTGTGCGCGCTGT 219  
203 AGTGTGCACTGACCTGGGCTCAGAGCTTCCACTTGCAGCTTTCTTCCACAGGTGC 262  
220 GGTGTGCTGCTGGGAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
263 ATCCCTGAAGAGCTGTGTGCGAGGTTGTGAGAGACTCTGCGAGGGCAACAGAGAA 322  
280 CTGCTGAAGAGCTGTGTGCGCGCGCTGTGAGAGCTGTGAGAGCGCGCGCGCG 339  
323 CTGCTGCTTTTGGCTTGTAGCTGTAAAGCGCGAGGCGCGCGCTTCCATGGCTT 382  
340 CTGCTGCGCTTGGCTTGTGCTGTGAGCGGCGCGCGCGCGCGCGCGCGCGCTT 399  
383 CACTAGTAGCTGCTACTCTTGTGCGCAACACTGTTATTGAGACCTGCGTGTGAGTG 442  
400 CACCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459  
443 TGCATGATGCTACTGTTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 502  
460 GCGGTGGGGCTGTGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 519  
503 CTGTGCTCTTTATCTTGTGTCGCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 562  
520 CTGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAAAATTTGTGCGCACCGGATATCTGCGCTCTGTGCTGCTGCTGCTGCT 622

Db Qy  
580 GTACCAGCTCGGCGTGCACCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639  
623 CCAGCCGCTGGGAGGAAATTTCACTAACTTAGTTCCTTAAAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGGCTGGAAACCATAG 672  
683 TCGCCAGGAGACACCGAAACCCCTTGCCATCTCGAGGTACAAAGAGGCATCTGAG 742  
673 CFTCAGGAGGCGGGGTCTCCCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCG 732  
743 TCTCACAGTACAGTGTGCTTACGTAGAGGCGGAGATGCTATCTGTCTCCGAGAT 802  
733 CAGTGCAGCCGAGTCTGCTGCTGCGGAGGCGCGCGCGCGCGCGCGCGCGCG 792  
803 GAGGAGGAGACCCACAGCAGGTGTCTACCAACCCCATCAGGCAATCATGGGTGCCAAG 862  
793 GAGCGGACCGCGCTTGGGAGGGTCTCGGCGCGCGCGCGCGCGCGCGCGCG 852  
863 TCTGTCTGGTCCCGCGAGGTGCTACT-----GCAGAGAAAGATTTGTCTTTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCTACCTGCCAGACCCCGGAGGAGCCACCTCTTGA 912  
914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGCGTGTGTGTCTGTAAACACAGCCAG 970  
913 GGTGTGCTCTCTGGCAGCGGCTCTCCACCCATCGGTGGCGCGCGCGCGCGCG 972  
971 CTCACAT---CTCTGTCTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1027  
973 CCCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGCGCTTAAACCCCTATT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCA---CAAGGAGAGCTGGGCGCTCT 1089  
1088 CTTACTCAGCAACCTCCAGCTAACTTGACTGGGCGCGAGAGACTGGTGGAGATCATCT 1147  
1090 CTTACTCAGCTCTCTCAGGCGCGCGCTGCTGCGTGTGGAGCTCGTGGAGCATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGACATCAGGACCACTCTGCGAGGACACACCGCTCTATCGGTG 1207  
1150 TCTGGGTTCCAGCGCTGTATGCCAGGACTCCCCGAGTTGGCGCGCGCTGCCCGAG 1209  
1208 ATACTGCAATCGCGCGCTGTTCACACAGCTGTGTGTGAAACCATGAGAGTGCATA 1267  
1210 CTACTGCAATCGCGCGCTGTCTGAGCTGTCTGGAACCAACGCGAGTGGCCCTA 1269  
1268 TGTCAACTCTCAGGTCAATTCAGGTTTGAACAGCAACCAACAGGTTGACAGATGC 1327  
1270 CGGGGTCTCTCAAGACGCACTGCGCGCTGCGAGTGTGGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACCGCGCG 1342  
1330 CTGTCCCGGAGAACCCCGAGGCTCTGTGGGCGCGCGCGCGCGCGCGCGCGCG 1389  
1343 ACCGACCTCATGATTTGCTCGCTGCGAGAGTCTCTGCGAGGTATATGTTTTCT 1402  
1390 CCGTGGCTGTGCTGCTGCTGCGCGAGCAGCGCGCGCGCGCGCGCGCGCGCG 1449  
1403 TCGGGCTGTCTCTGCAAGGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462  
1450 GCGGCGCTGCTGCGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509  
1463 CCGTCTCTTAAAGAACTTAAAGAACTTCTGCTTTGGGAAATACGGCAAGCTATCAT 1522  
1510 CCGTCTCTCAGGAACACCAAGAACTTCTCTCTGCGGAGAGCATGCGCAAGCTCTGCT 1569  
1523 GCAGAACTGATGTGAAGATGAAGTAGAGTTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
1570 GCAGAGCTGCTGCTGGAAGATGAGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCT 1629  
1583 GAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642  
1630 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689

1643 GTTCTGGCTGATGGACACATACGTTGGTACAGCTGCTTAGGTGATCTTTTACATPACAGA 1702  
1690 GCACCTGGCTGATGAGTGTAGCTGCTGAGCTGCTGAGTCTTTCTTTTATGTACGGA 1749  
1703 GAGCACATTCACGAAGAAACAGGCTCTTCTTACCGTAAAGAGTGTGGAGCAAGTGA 1762  
1750 GACCAGTTTCAAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTGGAGCAAGTGA 1809  
1763 GAGCATTTGGAGTACAGCAACACCTTGAGAGAGTGGCTACGGAGAGTGTCAACAGAGGA 1822  
1810 AAGCATTTGGAAATCAGACAGCACTTGAAGAGGTTGAGCTGCGGAGTGTGGAAGCAGA 1869  
1823 GGTACAGCATCACAGAGACACCTGGCTAGCATGCCCCATCTGCAGAGTGGCTTCATCCC 1882  
1870 GGTACAGCATCGGGAAGCCAGGCGCCGCTGTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTCGGCCCATTTGTGAACATAGTTATAGCATGGTACAGAGCTTT 1942  
1930 CAAGCCCTGAGCGGCTCGGCCGATTTGTGAACATAGTACTGCTGTGGAGGCCAAGCTT 1989  
1943 GGGCAGAAAGGAAGCAGGCGCCAGCATTTCAACCCAGCGTCTCAAGACTCTCTTTCAAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCCAGGCTCTCACTCGAGGCTGAAGCACTGTTCAAGGCTGT 2049  
2003 CAACATAGAGCGAGCAAAACATCTCTCACTTATGGGCTCTCTGATCGGTATGATGA 2062  
2050 CAACATAGAGCGGCGCGCGCCGCTCTGCGGCGCTCTCTGCTGGGCGCTGAGCGA 2109  
2063 CATCTACAGAGCTGCGGCGCTCTTGTGCTGCTGCTGCTCTGAGCAGACACCCAG 2122  
2110 TATCCACAGGCTCGCGCACCTTGTGCTGCTGCTGCGGCCAGGACCGCGCGCTGA 2169  
2123 GATGTACTTTTAAAGCAGATGTGACCGGCGCTATGATGCGATCCGCCAGCGTAAAGCT 2182  
2170 GCTGTACTTTTGAAGGTGATGTGAGCGGCGCTGACGACACCATCTCCCGAGCAGGCT 2229  
2183 GGTGGAGGTTTGTCCAAATATGATCAGGCACTCGGAGAGCAGCTACTGTATTCGCCAGTA 2242  
2230 CACGAGGTCATCGCCAGCATCATCA--AACCCAGAAACAGTACTGCGTGGTGGTA 2286  
2243 TGCAAGTGTGAGAGATAGCAGAGCCAGTCCAAAGTCTTTTAGAGACAGGTCAAC 2302  
2287 TGCGGTGATCAGAAAGCGCCCATGCGGCGCTGCGCAAGGCTTTCAGAGCCACAGTCTC 2346  
2303 CACCTCTCTGACCTCAGCCATACATGAGGCGAGTTCCTTAAAGCATCTGCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGCATCAGGAACCTCGGTTGTATCAGCAGAGCATCTCTATGATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATGCGGTGCTATCAGCAGAGCTCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCTCTTTGACTCTTCTGCTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTCCAGCTCTTCTAGCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGTGTCTATACGAGTCCAGGCACTCCCGAGGCTCCAGGCTTCCACCTGCTGCTGCTG 2542  
2521 CAGTCTCTAGCTCAGTGGCAGGGATCCCGAGGCTCCATCTCTCCAGCTGCTGCTG 2580  
2543 CAGTCTCTTTTGGAGATGAGAAACAGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2602  
2581 CAGCCTGTGCTAGCGGACATGAGAAACAGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2662  
2641 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
2663 CTTCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2722  
2701 CTTCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760

QY 2723 GACAGTGTGAACCTCCCTGTGAGAGCTGTACCTGGTGGTGCAGCTCCATACACAGCT 2782  
DB 2761 GACAGTGTGAACCTCCCTGTGAGAGAGAGAGCCCTGGTGGACGGCTTTTGTTCAGAT 2820  
QY 2783 GCCTGTCTCACTGCTGCTGCTTCCCTGGTGTGGTGTGCTGCGACACTCAGACTTTGGAGGT 2842  
DB 2821 GCCGGCCACAGGCTATTCCTGCTGGTGGGCTGCTGCTGATACCCGACCTCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACCTTCCAGAG 2902  
DB 2881 GCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAAGCTGGGAAGACCATCGGAACAAGCTCCTGCTGCTGGTGTGAGAGT 2962  
DB 2941 CGGCTTCAAAGCTGGGAAGAACATGCTCGCAAACTCTTTGGGGCTTGGCGCTGAAGGT 3000  
QY 2963 TCAGGCTCTATTTCTAGACTTCAGGTGAACAGCTCCAGACCTCAGACAGTCTGCATCAATATA 3022  
DB 3001 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCTCCAGAGCTGCTGCGACCAATCTA 3060  
QY 3023 CAAGATCTTCTGCTTCAAGCTTACAGCTTCCATGATGTGATTCAGCTTCCCTTTGA 3082  
DB 3061 CAAGATCTTCTGCTGCTGAGCTTACAGCTTTCAGCATGTGCTGCTGAGCTCCCATTTCA 3120  
QY 3083 CCAGCGTGTAGAGAGAACCTCACATTTCTTGGGATCATCTCCAGCCAAAGCATCTG 3142  
DB 3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTGGGCTCATCTCTGACAGGCTCTCCT 3180  
QY 3143 CTGCTATGCTTCTTCAAGGTCAAGAATCCAGGAATGACATNAAG-----GCTC 3193  
DB 3181 CTGCTACTCCATCTGAAAGCCAGAAAGCCAGGATGTGCTGGGGCCCAAGGGCGCGC 3240  
QY 3194 TGCTCTTCTTCTGCTGAGCGCACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3253  
DB 3241 CGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
QY 3254 GGTGCTCATTTCTGCTCATCTCAAAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3313  
DB 3301 GACTGACACCGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
QY 3314 GCTGTGCGGAAGCTCCAGAGCGGACATGACCATCTTAAAGCTGAGCTGAGCTGAGCTGAGCTG 3373  
DB 3361 GCTGAGTGGAGCTCCCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
QY 3374 CCTAAGCAGACAGCTTTCAGACCATTTTGGACTAA 3407  
DB 3421 ACTGCTCTCAGACTTCAAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3454

## RESULT 8

AR182221  
LOCUS AR182221 4015 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 1 from patent US 6337200.

ACCESSION AR182221

VERSION AR182221.1 GI:20225137

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE

1 (bases 1 to 4015)

AUTHORS

Morin, G.B.

TITLE

Human telomerase catalytic subunit variants

JOURNAL

Patent: US 6337200-A 1 08-JAN-2002;

Location/Qualifiers

FEATURES

1..4015

source

/organism="unknown"

BASE COUNT 663 a 1363 c 1275 g 714 t

ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGCCCTTGAGCACAAATGACCCCGCGCTCTCTGGTGGCCCCCGGTGGCTCTCTGTGCGC 82

40 CCCGCCCACCCCGCATGACCGCGCGCTCCCGGCTCCGAGCCGTGTGGCTCTCTGTGCTGCG 99

83 CAGCCGATACCGGAGGTGTGGCCGTGCAACCTTTGTGGGGCGCTGTGGGCCCGGAGGG 142

100 CAGCACCTACCGGAGGTGTGCCGTGTGGCACCTTGTGCGGCGCTGTGGGCCCCGAGGG 159

143 CAGCGGCTTTGTGCACCCCGGGACCCGAGATCTACCGCACTTTTGGTTGCCCAATGCCT 202

160 CTGCGCGTGTGTGACGGCGCGGACCCGCGCGCTTTCCGCGCGCTGGTGGCCAGTGCCT 219

203 AGTGTGCATCAGCTGGGGCTCAGAGCTCCACTCCGACCTTTCTCCCTTCACACGAGTGC 262

220 GGTGTGGTGGCCCTGGGACGACGGCGCGCCCCCGCGCGCCCTCTTTCCGCGAGGTGTC 279

263 ATCCCTGAAAGAGCTGTGGCGCAGGGTTGTGAGAGACTCTCGAGAGCGCAACGAGAGAAA 322

280 CTGCTGAAGGAGCTGTGGCCCGAGTGTGAGAGGCTGTGCGAGCGCGCGCGCAAGAA 339

323 CGTGTGGCTTTGGCTTTGAGCTGTTAACAGAGCCAGAGCGCGGCTCTCCATGGGCTTT 382

340 CGTGTGGCTTTGGCTTTGAGCTGTGAACGGGCGCGCGCGGGCGCCCCCGCGAGGCTTT 399

383 CACTAGTAGAGCTGGTGTAGCTACTTGGCCCAACTGTTATTGAGACCTTGGCTGTGAGTGG 442

400 CACCACAGCGTGGCAGCTACTCTGCCCCAACAGCGTGACCGACGCACTGCGGGGAGCGG 459

443 TGCATGAGATGCTACTGTTGAGCGAGTGTGGCGACGACCTGTGTGTTACTGTGTGGCACA 502

460 GCGGTGGGGGTGTGTGTGGCGCGGTGTGGCGACGACGTGTGTTCACTGTGTGGGACG 519

503 CTGTGCTCTTTATCTTCTGTGTGCCCGCAGCTGTGCTACCAAGGTGTGTGGTCTCTCCCT 562

520 CTGTGCGCTCTTTGTGTGTGTGGTTCGACGTGTGCGCTTACCAAGTGTGCGGCGCGCGCT 579

563 GTACCAATTTGTGCCACCAAGATATGTGGCCCTCTGTGTCCGTACTGTACAGGCCCAAC 622

580 GTACAGCTCGGCGCTGCCACTCAGGCGCGGCGCGCGCCACACACGCTAGTGACCCCAAG 639

623 CCGACCGGTGGCAGGAATTTCACTAACCTTAGTGTCTTACACAGATCAGAGCAGTAG 682

640 GCGTGTGGATGTC-----GAAACGGGCTCGGAACCATAG 672

683 TCGCCAGGAAGCACCGAAACCCCTTGCCCTTCGAGGTACAAAGAGGCATCTGAG 742

673 CGTCAGGAGGACCGGGTCCCTCGGCGCTGCCAGCCCGGTGCGAGAGGCGCGGGG 732

743 TCTCACAGTACAAAGTGTGCTCAGTTAAGAGGCGAGATGCTATCTGTCTCCGAGAGT 802

733 CAGTGTCCACGCCGAAGTCTGCGGTGTGCCAAGAGCGCCAGCGGTGGCGCTGCGCTGAGCC 792

803 GGAGAGGAGGACCCACAGGAGGTGTCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862

793 GGAGCGGAGCGCGTGTGGCAGGGGTCTTGGGCCCATCCGCGGACAGGACGCTGTGACCGAG 852

863 TCCTGCTCGGTCCCCCGAGTGCCTACT-----GCAGAGAAAGATTTGTCTTTCTAA 913

853 TGACCGTGGTTCTGTGTGTGTGTCACTGTCCAGACCGCGCGGAAGCGCACTCTTTGGA 912

914 AGGAAGAGGTGTCACTGAGTCTCTC---TGGGTGGGTGTCTGTCTTAACACAGAGCCGAG 970

913 GGGTGTGCTCTGTGGACGGCGCACTCCACCCATCCGTGGGCGCGCACAGCACCGCGGG 972

971 CTCCACAT---CTCTGTGTGCACCCCGGCCAAAATGCTTTTCACTCAGGCCCATTTAT 1027

973 CCCCCATCCACATCGCGGCCACCACTGCTCCGACACCGCTTGTCTCCCGGTGACGC 1032

1028 TGAGACACAGACATTTCTTCTTACTTCAGGGGAGATGGCCAGAGCGCTCTAAACCCCTCAT 1087

1033 CGAGACCAAGACATTCCTCTACTCTCTAGCGCA---CAGAGGAGAGCTGTGCGGCTCTCT 1089

1088 CCTACTCAGCAACCTCCAGCGCTAACTTGACTGGGGCGAGAGACTGGTGGAGATCATCTTT 1147

Db	1090	CTCTACTCAGTCTCTGAGGCCAGCCTGACTGCGCTCGAGGCTCTGAGACCAATCTT	1149
Qy	1148	TCTGGGCTCAAGSCCTAGGACATCAGGACCACTCTGCGAGGACACACCGTCTATCGCGTCG	1207
Db	1150	TCTGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCCAGCG	1209
Qy	1208	ATACTGGCAGATGCGGCCCTCTGTTCCAAACAGCTGCTGGTGAACCATGACAGATGCCAATA	1267
Db	1210	CTACTGGCAATGCGGCCCTGTTTCTGGAGCTGCTTGGGAACCAACGCGCAGTGCCTTA	1269
Qy	1268	TGTCAGACTCTCTCAGGTTCATTTGACAGTTCGAAACAGCAACAAACAGGTGACAGATGC	1327
Db	1270	CGGGTGCTCTCAAGACGACTGCCCCGTGCGAGCTGGGTCACCACAGCAGCGGTGT	1329
Qy	1328	CT-----TGAAACACAGCCC	1342
Db	1330	CTGTGCCCGGAGAAGCCCAAGGGCTCTGTGGCGGCCCGGAGGAGGAGCACAGACCC	1389
Qy	1343	ACCGCACCTCATGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGATATATGTTTCT	1402
Db	1390	CGCTCGCCTGGTGAGCTGCTCCGCCAGCACAGACGCCCTGGCAGGTGTACGGCTTCGT	1449
Qy	1403	TCGGGCTGTCTCTGCAAGGTGGTGTCTGCTAGTCTCTGGGTACCAAGGCACAAATGAGCG	1462
Db	1450	CGGGCTGTGCTGCGCGGCTGTGTGCCCGCAGCGCTCTGGGCTCCAGGCACAACGAACG	1509
Qy	1463	CGCTCTTTTAAGAACTTAAGAAAGTTCACTCTGTGGGGAATAACGCGACGCTATCACT	1522
Db	1510	CGCTTCTCTCAGGAACACCAGAAAGTTCACTCTCCCTGGGGAAGCAATGCCAAGCTCTCGCT	1569
Qy	1523	GCAGAACTGATGTGAAGATGAAGTAGAGGATGCCCACTGGCTCCGACAGCCCGGG	1582
Db	1570	GCAGAGCTGACGTGGAAGATGAGCTGCGGACTGCGCTTGGCTGCGCAGAGSCCCAGG	1629
Qy	1583	GAAGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCTCGGTACGTTCTCT	1642
Db	1630	GGTTGGCTGTCTCCGCGCAGCAGCACCGTCTGGCTGAGGAGATCTCGGCCAAGTTCTCT	1689
Qy	1643	GTTCGGCTGATGACACATACGTGTGACAGCTGCTTAGTTCATTCTTTATCATCACAGA	1702
Db	1690	GCATGGCTGATGAGTGTGTAAGTCTGTCAGAGTCTGTCAGGTCCTTTATGTACGGA	1749
Qy	1703	GAGCACATTCAGAAAGAACAGGCTCTTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA	1762
Db	1750	GACCACGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGACCAAGTTGCA	1809
Qy	1763	GAGCATTGGGTGAGGCAACACCTTGAGAGAGTGGGCTACGGAGCTGTCCACAGAGGA	1822
Db	1810	AAGCATTGGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTGTCCGAAGCAGA	1869
Qy	1823	GGTCAGGCATCACAGGACACCTGGCTAGCCATGCCATCTGCGACTCGGCTTCATCCC	1882
Db	1870	GGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGTCTGCTCAGCTCCAGCTTCATCCC	1929
Qy	1883	CAAGCCCAACGGCTCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT	1942
Db	1930	CAAGCTCAACGGGCTGGGCCGATTTGTGAACATGACCTACGTCGTGGAGCCAGAACTT	1989
Qy	1943	GGCGAAGGAAGCAGGCCCCAGCATTTCAACCAGCGGTCTCAAGACTCTCTTCAGCATGCT	2002
Db	1990	CGCAGAGAAAGAGGSCCAGCGCTCTACCTCAGGGTGAAGGCATGTTTCAGGCTGCT	2049
Qy	2003	CAACTATGAGCGGCAAAAACATCCTCACTTATGGGCTCTCTGTATCGGGTATGAATGA	2062
Db	2050	CAACTACAGCGGGCGGGCGCCCTCTCTGGCGCTCTCTGTGCTGGGCTTGGACGA	2109
Qy	2063	CATCTACAGGACCTGGCGGCTTTGTGTGGTGTGGGTCTCTGAGACAGACACCCAG	2122
Db	2110	TATCCACAGGCTTGGCGCACTTGTGTGTGTGGGCCCAAGNACCGCCGCTGA	2169
Qy	2123	GATGTACTTTTAAAGCAGATGACCGGGCCCTATGATGCCATGCCCAAGGTAAGCT	2182



2170 GCTGTACTTTGTCAAGTGGATGTGACGGGCGGTACAGACACCATATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTGTTGCCAATATGATCAGGACCTCGAGAGACAGTATGTATPCCGCCAGTA 2242  
2230 CACGGAGGTCAATGCCAGCATCATCA--AACCCAGAAACAGTACTGCGTGCCTCGGTA 2286  
2243 TGCAATGTTGCGGAGATAGCAAGGCCAAGTCCACAAGTCTCTTTAGAGACAGGTCAC 2302  
2287 TGCGTGTTCAGAGGCCCCCATGSGGACGCTCGGAGGCTTCAGAGCCACAGTCTC 2346  
2303 CACCTCTCTGACTCCAGCCATACATGAGGCCAGTTCTTAAAGCATCTGCAAGATTCAGA 2362  
2347 TACCTTGACACAGCTCCAGCCGTACATGCGACAGTTCTGCTGCTCACTGTCAG-----GA 2400  
2363 TGCAGTGTCACTGAGGAATCCGTTCTCATCGAGCAGACGATCTCTATGAATGAGACAG 2422  
2401 GACAGGCCCGCTGAGGATGCGCTGCTCATCGAGCAGAGCTCTCTCTCATGAGTGGCCAG 2460  
2423 CAGCAGCCTGTTTGCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482  
2461 CAGTGGGCTCTTTCAGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGTGTCTATACGAGGTCAGGAGGCTATCCCGAGGCTCCAGGCTATCCAGCTCTGCTG 2542  
2521 CAGTCTCTACGTCAGTGCAGGAGGATCCCGAGGCTCCATCTCTCTCATGCTGCTG 2580  
2543 CAGTCTGTTTTCGAGACATGAGAACAAAGCTGTTTGTGAGGTCGAGCGGAGTGGTT 2602  
2581 CAGCTGTGCTACGCGCACATGAGAACAAAGCTGTTTGGGGGATTCGCGCGGACGGCT 2640  
2603 GCTTTTACGTTTCTGATCATCTTCTGTTGTTGAGCTCTCATCTTGAGCAAGCAAAAC 2662  
2641 GCTCTGCTTGTGATGATTTCTGTTGTTGATGACACCTCATCTCACCGCGGAAAC 2700  
2663 CTCTCTCAGACCTGCTCATGCGGTTCTGATGATGAGTATGAGTATGATGATGATGATGAT 2722  
2701 CTCTCTCAGACCTGCTCATGCGGTTCTGATGATGAGTATGAGTATGATGATGATGATGAT 2760  
2723 GACAGTGTGATCTCTCTGAGGCTGCTGATCTCTGAGTGTGAGTGTGAGTGTGAGTGT 2782  
2761 GACAGTGTGATCTCTCTGAGGCTGCTGATCTCTGAGTGTGAGTGTGAGTGTGAGTGT 2820  
2783 GCCTGCTCATGCTCTCTCTGAGGCTGCTGATCTCTGAGTGTGAGTGTGAGTGTGAGTGT 2842  
2821 GCGGCGCACGCGCTATCTCTCTGAGGCTGCTGATCTCTGAGTGTGAGTGTGAGTGTGAG 2880  
2843 GTTCTGACTACTCAGGTTATGCGGACCTCAATTAAGACGAGCTCACTCTCCAGAG 2902  
2881 GCAGAGGACTACTCCAGCTATGCCGACCTCTCATCAGAGCAGTCTCATCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGAGACCATGCGAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGGAGAACATGCGTCCGAACTCTTTCGCGGCTCTGCGGCTGAGGT 3000  
2963 TCAGGCTTATTTAGAGTTCAGGTTGAGGCTTCCAGAGCTCTGATCAATATATA 3022  
3001 TCAGAGCTGTTTCTGATTTGAGTTCAGGTTGAGGCTTCCAGAGCTCTGATCAATATA 3060  
3023 CAGATCTTCTGCTTCAAGCTTCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 3082  
3061 CAAGATCTCTCTGCTGAGGCTTCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 3120  
3083 CCAGGCTTATAGGAAGCTTCAATTTCTTGGGCTATCTCTCAGGCTTCAATCTGATCAGT 3142  
3121 TCAGCAAGTTTGAAGAACCCCATTTTTCTGCGGCTCATCTCTGACACGCGCTCTCT 3180  
3143 CTGCTATCTATCTGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 3193  
3181 CTGCTATCTTCTGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 3240  
3194 TGCTCTCTTCTCTGAGGCTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
3241 CGGCTCTCTGCTCTCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300

QY 3254 GGTGCTCATTTCTGTCAATCTACAAATGCTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
Db 3301 GACTCGACACCGTGTACCTAGCTGCTCACTCTCTGGGTCCTCTGAGGACAGCCGAGCGCA 3360  
QY 3314 GCTGTGCGGAGAGCTCCAGAGCGGACCAATGACCACTTAAAGCTGAGTGCAGTGCAGCCAGC 3373  
Db 3361 GCTGAGTGGAGAGCTCCCGGAGCGAGCTGACTGCTCTGGAGGCGGAGCAACCCGCGC 3420  
QY 3374 CTTAAGCAGACAGCTTTTCAAGCACTTTTGGACTAA 3407  
Db 3421 ACTGCTCTGAGCTTCAAGACCATCTGAGCTGA 3454

RESULT 9  
AR224455  
LOCUS AR224455 4015 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 1 from patent US 6440735.  
ACCESSION AR224455  
VERSION AR224455.1 GI:23333293  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Gaeta,F.C.A.  
TITLE Dendritic cell vaccine containing telomerase reverse transcriptase for the treatment of cancer  
JOURNAL Patent: US 6440735-A 1 27-AUG-2002;  
FEATURES Location/Qualifiers  
source 1..4015  
/organism="unknown"  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 58.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGGCTTGTGAGCACATGACCGCGCTCTCTGTTGCCCGGCTGCTCTCTGCTGCG 82  
Db 40 CCGGCTTGTGAGCACATGACCGCGCTCTCTGTTGCCCGGCTGCTCTCTGCTGCG 99  
QY 83 CAGCGATACCGGAGGTGCGCGCTGCGAACCTTTGTGCGGCGCTTGGGCGCCGAGG 142  
Db 100 CAGCGATACCGGAGGTGCGCGCTGCGAACCTTTGTGCGGCGCTTGGGCGCCGAGG 159  
QY 143 CAGCGCTTGTGCAACCGCGGAGCCGAGATCTACCGCACTTTGGTGGCCCAATGCT 202  
Db 150 CTGGGCTGTTGAGCGCGGAGCCCGCGCTTTCGCGCTGTTGGCCAGTGCCT 219  
QY 203 AGTGTGATGACTGGGCTCAGCGCTCAGCTCCAGTCTTCTTCCAGCGAGTGC 262  
Db 220 GGTGTGCTGCTTGGGCTCAGCGCTCAGCTCCAGTCTTCTTCCAGCGAGTGC 279  
QY 263 ATCCCTCAAGAGCTGTTGCGGAGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTTGAAGAGCTGTTGCGGAGGTTGTGAGAGACTCTGCGAGCGCGCGGAGAA 339  
QY 323 GGTGTGCTGCTTGGGCTTGAAGTCTTAAAGAGCGAGAGCGGCGCTTCCATGGCTT 382  
Db 340 GGTGTGCTGCTTGGGCTTGAAGTCTTAAAGAGCGAGAGCGGCGCTTCCATGGCTT 399  
QY 383 CACTAGTAGGCTGCTGAGTCTTGTGCCCAACACTGTTATTGAGACCTTGGTGTGCTG 442  
Db 400 CACCAACGAGCTGCGAGCTTACCTGCCCAACAGGTGACCGCACTGCGGGGAGCG 459  
QY 443 TGCATGATGCTACTGTTGAGCGAGTGGGCGAGCACTGCTGCTGCTGCTGCTGCTGCT 502  
Db 460 GGTGTGCGGCTGCTGCTGCGCGCTGCGGCGAGCACTGCTGCTGCTGCTGCTGCTGCT 519  
QY 503 CTGTGCTCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562



520 CTGCGCGCTCTTTGTGCTGTGTGGCTCCACAGCTGCGCCTACACAGTGTGTGGGGCGCGCCGCT 579

563 GTACCAAAATTTGTGCCACACCGATATCTGGCCCTCTGTGTCCGTAGTTACAGGCCAC 622

580 GTACCAGCTGGCGCTGCCACTCAGGCCGGGCCCGCCACACGCTAGTGTGACCCGAG 639

623 CCGACCCGTGGCGAGGAATTCACATACTTAGTTCTTACAACAGATCAAGACAGTAG 682

640 GCGTGTGGATGC-----GAA CGGGCTCGGAACCATAG 672

683 TCGCCAGGAACACCGAAACCCCTGGCCTTGCCATCTCGAGGTACAAGAGGCATCTTGAG 742

673 CGTCAGGAGGCGCGGGTCCCTTGCGCTGCCAGCCCGGGTGCAGAGGCGCGGG 732

743 TCTCACAGTACAGTGTGCTTCAGTAAAGGCCAGATGCTATCTGTCTCCGAGGT 802

733 CAGTGCCAGCCGAAGTCTGCGCTGTGCCAAGAGGCCAGCGGTGGCGCTGCCCTGAGCC 752

803 GGAGAGGGAACCCACACAGGAGGTGTCAACAACCCCATCAGGCAAAATCATGGGTGCCAAG 862

793 GGAGCGACGCCGTGGGCGAGGGTCTTGGGCCCAACCGGGCAGGACGCGTGACCGAG 852

863 TCCTGCTCGTTCGCCCGAGGTGCTACT-----GCAGAAAGATTTGTCTTCTAA 913

853 TGACCGTGTGTTCTGTGTGTGTCACCTGCCAGACCCCGCAAGAACGACCTCTTTGGA 912

914 AGAAAGGTGTCACTGAGTCTCTC---TGSGTGGTGTGCTGTAAACACAGCCCGAG 970

913 GGGTGGCTCTTGGCACGGCACTCCACCACTCGTGGGCCGCGAGCACACGCGGG 972

971 CTCACAT---CTCTGTGTCAACACCCCGCAAAATGCTTTTCAGCTCAGGCCAATTTAT 1027

973 CCCCCATCCACATCGCGGCCACCACTGCTCGGACACGCGTTGCTCCCGGTGTACGC 1032

1028 TGAGACCAAGCATTTCTTTACTTCCAGGGAGATGSCCAAGCGCTCTAAACCCCTCATTT 1087

1033 CGAGACCAAGCATTTCTCTACTCTCTCAGCGCA---CAAGAGACAGCTGTGGGCCCTCTCTT 1089

1088 CTTACTCAGCAACTCCAGCCTCACTTTGACTGGGGCGAGGACTGTGTGGAGATCATCTTT 1147

1090 CTTACTCAGCTCTCTGAGGCCAGCGCTGACTGGCGTCTGGAGCTCGTGGAGACCATCTTT 1149

1148 TCTGGCTCAAGGCTTAGACATCAGGACCACTTCACAGGACACACCGTCTATCGGCTCG 1207

1150 TCTGGGTTCCAGGCCCTGGATGCGAGGAGCTCCCGCAGGTTGCCCGCCCTGCCCCACGC 1209

1208 ATACTGGCAGATCGGCCCTTTTCAAACAGCTGTGTGTGAACCATGCAGAGTCCCAATA 1267

1210 CTACTGGCAATGGGCCCTGTTTCTGAGAGCTGTGTGGAAACCAACGCGAGTGGCCCTTA 1269

1268 TGTGAGCTCTCTCAGGTCAATTCAGGTTTCGAAACAGCAACCAACAGGTGACAGATGC 1327

1270 CGGGGTGCTCTCTAAGACGCACTGCCCGTGTGCGAGCTGGCGGTCACTCCACAGCGCGGTGT 1329

1328 CT-----TGACACACGACCC 1342

1330 CTGTGCCGAGGAAGCCCGAGGCTGTGTGGGGCCCCCGAGAGAGAGACACAGACCC 1389

1343 ACCGACCTCATGGAATTTGCTCGCCCTGCACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402

1390 CCGTCCGCTGTGAGCTGTCTCGCCAGACACAGACGCCCTTGGCAGGTGTACGCTTCGT 1449

1403 TGGGCTGTCTCTGCAAGGTGTGTCTGTAGTCTCTGGGTACCGAGCAATGAGCG 1462

1450 GGGGCTGTCTGTGGCGGGCTGTGTGCCCCAGGCCCTCTGGGCTCCAGGCAACAGCAACG 1509

1463 CGCTTCTTTAAGAACTTAAAGAAAGTTTCATCTCGTTGGGGAATACGGCAAGCTATCACT 1522

1510 CGCTTCTCTAGAAACCAAGAGTTTCATCTCCCTGGGAGAGCATGCGAGCTCTCGCT 1569

1523 GAGAACTGATGTGGAGATGAAGTAGAGGATGCCACTGGCTCCCGACGACCGCGGG 1582

1570 GAGAGGTGAGTGTGAAGATGACGCTGTGGGATCTGCGCTGTGGTTCGCGAGGAGCCAGG 1629

Qy	1583	GAAGACCGTGTCCC	CGCTCAGACGACACCGTCTGAGGGAGAGGATCCTGGGTACGTTCTCT	1642
Db	1630	GGTTGGCTGTCTTCCGGCCG	CAGACACCGTCTGGGTGAGGAGATCCTGGCCAAATTCCT	1689
Qy	1643	GTTCGGCTGATGACACATAC	AGTGGTACAGCTGCTTAGGTCAATCTTTTATACATCACAGA	1702
Db	1690	GCATCTGGCTGATGAGTGTG	TACGTCGTGAGCTGCTCAGGCTCTTTCTTTATGTCACGGA	1749
Qy	1703	GAGCAATTCAGAAAGCAAGG	CTTTCTTACCGTAAAGTGTGTGGAGCAAGCTGCA	1762
Db	1750	GACCAAGTTTCAAAGAACAGG	CTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA	1809
Qy	1763	GAGCAATTCGAGTCAGGC	CAACCTTCAGAGAGTGGGGCTACGGGAGCTGTCAACAAGAGGA	1822
Db	1810	AGCATTTGGAAATCAGACAG	CACATTGAAGAGGGTGCAGCTGCGGGAGCTGTCGGAAGCAGA	1869
Qy	1823	GGTCAGGCATCACACAGACA	CTCTGGCTAGCCATGCCATCTCAGACTGCGCTTCATCCC	1882
Db	1870	GGTCAGGCAGATCGGAAAC	CCAGGCCCGCTCTGAGCTCAGACTCGCTTCATCCC	1929
Qy	1883	CAAGCCCAACGGCTCGGG	CGCCCATTTGTGAACATGAGTTATAGCATGGGTACACAGAGCTTT	1942
Db	1930	CAAGCTCGACGGCTGCGG	CCGATTTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTT	1989
Qy	1943	GGGAGAAGGACAGGCC	CCAGCATTTCCACCAGGGTCTCAAGACTCTCTTCAGCATGCT	2002
Db	1990	CGCGAGAAAGAGGGCC	CCAGCGTCTCACTCGAAGGTGAAGGCACCTGTTTCAGCGTGCT	2049
Qy	2003	CAACTATGAGCGGACAAAA	CAATCCTCACCTTATGGGGTCTTCTGTACTTGGGTATGAATGA	2062
Db	2050	CAACTACAGCGGCGCGG	CCCGCTCCTGGCGGCTCTGTGCTGGGCTGGACGA	2109
Qy	2063	CATCTACAGGACCTGGG	CGGCCCTTTGTGTGGGTGTGGTCTGTGGACACGACACCCAG	2122
Db	2110	TATCACAGGGCTTGGG	CGCACCTTCGTGTGTGGGGCCAGGACCCCGCCGCTGA	2169
Qy	2123	GATGTACTTTCTTAAGG	CAAGTGTACCGGGGCGCTATGATGCCATCCCCCAGGGTAAAGCT	2182
Db	2170	GCTGTACTTTGTCAAG	GTGGATGTGACGGGCGCGTACGACACCATCCCCCAGGACAGCT	2229
Qy	2183	GGTGGAGGTTTGGCAAT	ATGATCAGGCATCGGACCTCGGAGACGACTACTGTATCCGCCAGTA	2242
Db	2230	CACGAGGTCTATCCCA	GAGCATCA---AACCCAGAACAGTACTCGTGGCGTGGGTA	2286
Qy	2243	TGCAGTGTCCGAGAGAT	AGCAAGGCCAAAGTCCACAAGTCTTTTAGGAGACAGGTCAC	2302
Db	2287	TGCCGTGTTCAGAA	GGCCGCCCATCGGACGCTCCGCAAGGCTTCAAGAGCCACGCTC	2346
Qy	2303	CACCTCTCTGACCTC	CAGCATACATGGCCAGTTCCTTAGCATCTGCAAGGATTCAGA	2362
Db	2347	TACCTTGACAGACTT	CAGCCGCTACATGCAAGTTCGTGGCTACCTGTCAG-----GA	2400
Qy	2363	TGCCAGTGCATGAGGA	AACTCCGTGTCTATCGAGCAGGACATCTCTATGAATGAGAGCAG	2422
Db	2401	GACCAGCCGCTGAGG	GAATGCGGTCTATCGAGCAGAGCTCCTCCCTGAATGAGGCCAG	2460
Qy	2423	CAGCAGCGTGTTCAG	CTTCTCTGTCACATCTCCTGGTCACTAGTGTGCTGAAGATTTGTA	2482
Db	2461	CAGTGGCTCTTCAG	AGCTTCTCTAAGCTTCATGTGCCACACGCCGCTGCGCATCAGGGG	2520
Qy	2483	CAGGTGCTATACGAG	TGCCAGGCGATCCCCAGGGCTCCAGCCCTATCCACCTGCTCTG	2542
Db	2521	CAAGTCTTACGTC	AGTGCAGGGGATCCCGCAGGGCTCCATCTCTCACGGCTGCTCTG	2580
Qy	2543	CAGTGTGTTCGAG	ACATGGAACAGCTGTTGTGATGAGGTGCAGCGGATGGGTT	2602
Db	2581	CAGCTGTGTACG	CGGACATGAGAAACAGCTGTTTTCGGGGGATTCGCGGGACGGCT	2640
Qy	2603	GCTTTTACGTTTTGT	GATGACTTTCTGTTGGTGCGCCTCATCTTGGACCAAGCAAAAC	2662
Db	2641	GCTCTGGGTTTGGT	GATGATTTCTGTTGGTGACCTCACTCACTCACCCACGCGAAAC	2700



1028 TGAGACACAGACATTTCTTTACTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCATTT 1087  
1033 CGAGACCAAGACACTTCTTACTTCTCAGGCGA---CAAGGAGACAGCTGGGCGCTCTCTTT 1089  
1088 CCTACTCAGCAACCTCCAGCGCTAACTTTGACTGGGGCCAGGAGACTGGTGAGATCATCTT 1147  
1090 CCTACTCAGCTCTCTGAGGCCAGCTGACTGGGCTCGAGGCTGTTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGAACCACTCTGACGACACACCGTCTATCGGTCG 1207  
1150 TCTGGGTTCCAGGGCCCTGGATGSCCAGGGACTCCCGCAGGTTGGCCGCTGCCCAGCG 1209  
1208 ATACTGGCAGATGGGCCCTGTTTCAACAGCTGCTGGTGAACATGACAGAGTGCCAAATA 1267  
1210 CTACTGGCAATGCGGCCCTGTTTCTGAGAGTGTCTTGGGACACACGCGAGTGCCCCCTA 1269  
1268 TGTCAAGCTCTCAGGTCACANTCGAGTTTCGAAACAGCAAAACCAAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACGCACTGCCCGCTCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGCCC 1342  
1330 CTGTGCCGGGAGAGCCCCAGGCTCTGTGGGCCCCCGAGGAGGAGACACAGACCC 1389  
1343 ACCGCACCTCATGATTTGCTCCGCCCTGCACAGAGTCCCTGGCAGGTATATGGTTTTCT 1402  
1390 CGTCCGCTGTGTGACGTGCTCCGCCAGCAGCAGCGCCCTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGGCTGTCTGTGCAAGTGTGTCTGTACTCTGTGGGTACAGGCACAATGAGCG 1462  
1450 GCGGGCTGTCTGCGCGGCTGTGTGCCCCAGGCTCTGGGGTCCAGGACAAACGAAAG 1509  
1463 CGCTTTCTTAAAGAACTTAAAGAAAGTTTCACTCTGTGGGAAATACGCAAGTATCACT 1522  
1510 CGCTTCTCAGBAACACCAAGAAAGTTTCACTCTCCTGGGGAAGCNTGCCAAGTCTCGT 1569  
1523 GCAGGAACTGATGTGGAAGATGAAGTAGAGGATGCGCACTGCTCCGACAGAGCCCGGG 1582  
1570 GCAGAGCTGACGTGGAAGATGAGCGTGCAGGACTGCGCTTGGCTGCGCAGGAGCCAGG 1629  
1583 GAAGAACGCTGCTCCCGCTGCAGAGCAGCGTGTGAGGAGAGGATCTGGCTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGGCCGACAGACCGTCTCGGTGAGGAGATCTGGCAGTTCT 1689  
1643 GTTCTGGCTGATGACACATACGTGGTACAGCTGCTTAACTTCTTTTACATCAGAGA 1702  
1690 GCACCTGGCTGATGAGTGTGATCGTCTGAGCTGCTCAGGCTCTTTCTTTTATGTACGGA 1749  
1703 GAGCATTCCAGAGACAGGCTCTTCTTACCGTAAAGTGTGTGGAGCAGCTGCA 1762  
1750 GACCAAGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTGTGAGCAAGTTGCA 1809  
1763 GAGCAATTGAGTCAAGCAACCTTGAGAGGTGCGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCAATTGATCAGACAGCACTGAAGAGGTGCAGCTCGGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCAGCAGACACTGGCTAGCCATGCCCCATCTGACAGCTGGCTTCAATCCC 1882  
1870 GGTCAAGCAGCATCGGGAAGCAGGCGCCGCTGCTGAGCTCCAGATCCAGACTCCGCTTCAATCCC 1929  
1883 CAAGCCAAAGCCTGCGGCCCAATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAAGCTGACGGCTGCGGCCGATTTGTGAACATGAGTACTGTTGGGAGCCAGACGTT 1989  
1943 GGGCAGAAGGAGGAGCCAGCATTTCAACCAAGCTCTCAAGACTCTCTTCAAGCATGCT 2002  
1990 CCGCAGAGAAAAGAGGCGGAGCGTCTCACTCGAGGTTGAAGCACTGTTCAAGCTGCT 2049  
2003 CAACCTATGAGCGGCAAAACATCTCACTTATGGGTCTTCTGTACTGGTATGATGA 2062  
2050 CAACCTACGAGCGGCGCGGCCCTCTGGGCGCTCTGTGCTGGGCTGGAACA 2109

QY 2063 CATCTACAGGACCTGGCGGCGCTTTGTGCTGCGTGTGCGTCTTGGACACAGACACCAG 2122  
Db 2110 TATCCACAGGGCTGGCGCACCTTCTGCTGCTGTGGGGCCAGGACCCGCGCCTGGA 2169  
QY 2123 GATGTACTTTTAAAGCAGATGTGACCGGGCTTATGATGCAATCCGCCAGGGTAAGCT 2182  
Db 2170 GCTGTACTTTGTCAAGGTGATGTGACGGGCGGTACGACACCAATCCGCCAGGACAGCT 2229  
QY 2183 GGTGGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCACTGATCTGTATCCGCCAGTA 2242  
Db 2230 CACGGAGGTCAATCGCCAGCATCATCA---AACCCAGAAACAGTACTGCGTGTGCGTA 2286  
QY 2243 TGCAAGTGGTCCGAGAGATAGCAAGGCCAAAGTCCACAAGTCTCTTATGAGACAGGTCAAC 2302  
Db 2287 TGCGTGGTCCAGAAAGCGCCCATGGGCACTCGGCAAGGCTTCAGAGGCCAGCTCTC 2346  
QY 2303 CAGCTCTCTGAACCTCAGCAGCATACATGGGCGAGTTCTTAAAGCATCTGAGAGATTCTGA 2362  
Db 2347 TACCTTGACAGACCTCAGCGCTTACATGCGACAGTTCTGTGGCTCACCTGCGAG-----GA 2400  
QY 2363 TGCCAGTGCATCAGGAACTCCGTTGTCATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
Db 2401 GACAGGCCCTGAGGAGTGCCTGTCATCGAGCAGAGCTCTCTCTGAATGAGGCCAG 2460  
QY 2423 CAGCAGCTGTTTGACTTCTTCTGCACTTCTGCGTCAACAGTGTCTGTAAGATTTGGTGA 2482  
Db 2461 CAGTGGCTCTTTCGACGTCTTCTTACGCTTCTATGTCACACGCGGTGCGCATCAGGGG 2520  
QY 2483 CAGGTGCTATACGAGTGCAGGGGATCCCGAGGGCTCCAGCTTCCACCTCTGCTCTG 2542  
Db 2521 CAAGTCTCTAGCTCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
QY 2543 CAGTCTGTGTTTCCGAGACATGAGAAACAGCTGTTTGTGAGGTGCAGCGGATGCGGTT 2602  
Db 2581 CAGCTGTGCTACGCGCAGCATGAGAAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
QY 2603 GCTTTTACGTTTGTGATGAATTTCTGTTGTGAGCGCTCACTTGGAGCAAGCAAAAC 2662  
Db 2641 GCTCTGCGTTTGTGATGATTTCTTGTGTGAGACCTCACTCACTCCACCGCGAAAC 2700  
QY 2663 CTCTCTCAGCACCTGCTCCATGGCGCTTCTGATGATGGTGCATGATAAACTTCAGAA 2722  
Db 2701 CTCTCTCAGAACCTGCTGGTCCGAGGTGTCTCTGATATGGTCTGCTGGTGAATTCGGGAA 2760  
QY 2723 GACAGTGGTGAATTTCTCTGTGAGGCTTGGTACCTGTGGTGTGAGTCTCATACAGCT 2782  
Db 2761 GACAGTGGTGAATTTCTCTGTAGAGCAGGCGCTTGGTGGCAGCGCTTTTGTTCAGAT 2820  
QY 2783 GCCTGTCTACCTGCTCTTCTCTGCTGCTGGTGTGCTGCTGGACACTCAGACTTTTGGAGGT 2842  
Db 2821 GCGGCCCCAGGCTTATTTCCCTGTGGTGGCGCTGTCTGTGGATACCGGACCCCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAG 2902  
Db 2881 GCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCAGTCTCACTTCAACCG 2940  
QY 2903 TGTCTTCAAGCTGGGAAGACCATGCGGAACAGCTCTGCTGCTGCTTGGCGTTGAAGTG 2962  
Db 2941 CGGCTTCAAGGCTGGAGGAACATGGTTCGCAAACTCTTGGGGTCTTGGGCTGAAAGTG 3000  
QY 2963 TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAAGATCTCTCTGCTTACGGCTTACAGTTTCATGTGATGTGTGATTCAGCTTCCCTTGA 3082  
Db 3061 CAAGATCTCTCTGCTCAGCGCTTACAGGTTTTCAGCATGTGTGCTGCGAGCTCCCAATTTCA 3120  
QY 3083 CCAGCGTGTGTAGGAAGACCTCACAATTTCTTCTGGGCACTCATCTCCAGCAAGCATCTG 3142  
Db 3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTGCGGTCTCTCTGTGACAGGCTCTCCT 3180  
QY 3143 CTGCTATGCTATCTCTAGGCTCAAGAAATCCAGGAATGACACTAAAG-----GCCTC 3193



1463 CCGCTTCTTTAAGAACTTAAGAGTTTCACTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCTAGGAACACCAAGAGTTTCACTCTGTTGGGGAAGCATGCGCAAGCTCTCGCT 1569  
1523 GCAGGAACGATGTGGAGAGTAAAGTATAGAGATTCCTGCTCCGCGAGAGCCCGGG 1582  
1570 GCAGGAGCTGATGTGGAGAGTAAAGTATAGAGATTCCTGCTCCGCGAGAGCCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTCCAGAGCAACCGTCTGAGGAGAGATCTGGTTCCT 1642  
1630 GGTGGCTGTGTTCCGCGCCAGAGACACCGTCTGCTGAGGAGATCTGGCCAAAGTCTCT 1689  
1643 GTTCTGCTGATGAGACACATACGTTGTACAGTCTTAGTGTATCTTTTATCATACACAGA 1702  
1690 GCACTGGCTGATGAGTGTACGTGTGAGCTGTCTCAGGTCTTTCTTTTATGTCCACGA 1749  
1703 GAGCACATTCAGAGAAACAGGCTCTTTCTTACCGTAAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACCACTTTCAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGGAGTACGGCAACACCTTGAGAGAGTGGCTACGGGAGCTGTACAGAGAGA 1822  
1810 AAGCATTTGGAATCAGACAGACCTTGAAGAGGTTGCACTGCGGGAGCTGTGCGAAGCAGA 1869  
1823 GGTCAAGCATCACAGAGACCTGGCTTAGCCATGCCCCATCTGACAGACTGCGCTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGGAAGCAGGCGCCGCTGCTGAGCTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTGCGGCCCATTTGTGAACATGAGTTATAGCATGGTACAGAGCTTTT 1942  
1930 CAAGCCTGACGGGCTGCGGCCGATTTGTGAACATGAGTCTACGCTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAAGCAGGCGCCAGACATTTACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCCAGAGTCTACCTCGAGGTTGAGGCACTGTTTCAGCGTGTCT 2049  
2003 CAATATGAGCGGCAAAAATCTCTCACTTATGCGGCTTCTGTAAGTGGGTATGAATGA 2062  
2050 CAATACAGCGGCGCGCGCCCGCTCTGGGCGCTCTGTGCTGGGCGCTGGAAGA 2109  
2063 CATCTACAGAGCTGCGGCGCTTGTGCTGCGTGTGCGTGTCTGAGCCACAGACACCCAG 2122  
2110 TATCCAGAGGCTTGGCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169  
2123 GATGACTTTTAAAGCAGATGTGACGGGCGCTATGATGCCATCCCCCGAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCCCGAGCAGGCT 2229  
2183 GGTGGAGTTGTTCCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCGAGTA 2242  
2230 CACGGAGGTCAATGCGCCAGCATCA---AACCCCAAGACAGTACTGCTGCTGCTGCTG 2286  
2243 TGCACTGCTCGGAGATAGCAAGGCCAAGTCCCAAGTCTCTTTCAGGAGACAGTCAAC 2302  
2287 TGCGGTGGTCCAGAGGCGCCCATGGGACGTCGCAAGGCTTTCAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGAGGCGAGTTCCTTAAAGCATCTCGAGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACTGCAG---GA 2400  
2363 TGCCAGTGCATGAGGAACCTCGTTGTATCGAGAGAGCATCTCTATCAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGAGATGCGTGTGATCGAGAGAGCTCTCTCTCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGAATCTTCTCTGCACTTCTGCTGCTCAGTGTCTGTAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTCGACGCTTCTCTACGCTTTCATGTGCCACACCGCGTGGCATCAGGGG 2520  
2483 CAGTGTCTATACGAGTGGCAGGCGCATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCAGTGCAGGGGATCCCGCAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGGAGAACAAAGCTGTTTGTGAGTGCAGCGGAGTGGTT 2602

2581 CAGCCTGTCTACGGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGGACGGGCT 2640  
2603 GCTTTTACCTTTTGTGTGATGACTTCTGTTGGTGAACGCTCAGTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCGCTTGTGGTGGATGATTTCTTGTGGTGAACCTCAGCTCACCACCGGAAAC 2700  
2663 CTTCCTCAGACACCTGCTGCTCCATGGGCTTCTGAGTATGGGTGATGATAAACTTGCAGAA 2722  
2701 CTTCCTCAGGACCTGCTGCTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTTGCAGAA 2760  
2723 GACAGTGTGAACTTCTCTGCTGAGGCTGCTACCTGCGGTGCTGAGCTCCATACCACT 2782  
2761 GACAGTGTGAACTTCTCTGCTGAGGACGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCACGCGCTATTCCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGTACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCAGCTTCCAGAG 2902  
2881 GACAGCGAATCTCAGCTATGCCCGGACCTCTCATGAGCGCAGTCTCAGCTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAAGACCATCGGAACAAGCTCCTGTCGCTTGTGCGGTTGAAGTG 2962  
2941 CGGCTTCAAAGCTGGGAAGACCATCGCTCGCAAACTCTTGGGGTCTTGGCGCTGAAGTG 3000  
2963 TCAGGCTTATTTCTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCTGCTCAATATATA 3022  
3001 TCAGGCTTATTTCTGATTTGCAATGCAAGCTCCAGACGCTCCAGACGCTGCTGCTCAACATCTA 3060  
3023 CAAGATCTTCTGCTTCAAGGCTCAGGCTTACAGGTTCCATGCTGCTGCTGCTGCTGCTGCT 3082  
3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3083 CAGGCTGTAGGAGAGCTCAGATTTCTTCTGCGCATCTCTCCAGCCAGCATCTGCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCCACTTTCTGCGCTCATCTCTGCGCTCATCTCTGACCGGCTTCCCT 3180  
3143 CTGCTATGCTATCTCTGAGGTCAGAAATCCAGGAATGACACTAAAG-----GCCTC 3193  
3181 CTGCTACTCATCTCTGAAGCCAGAACCGCAGGATGTCGCTGGGGGCCAAGGGCGCGC 3240  
3194 TGCTCTCTTCTCTGAAAGCCGACATTTGGCTGTGCTACAGGCTTCTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTCTCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGTGCTCATTTCTGCTCATCAAAATGCTCTCTGCGACTCTGAGGACAGCCCAAAACT 3313  
3301 GACTCGACACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTCCGGAAGCTCCAGAGGCGACATGACCATCTTAAAGCTGAGTGCAGTCCAGC 3373  
3361 GCTGAGTCGGAAGCTCCCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CTTAAGCAGAGCTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCGCTCAGACTTCAAGACCATCTGAGACTGA 3454

## RESULT 12

AR263555

LOCUS

DEFINITION

AR263555

AR263555.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AR263555 Sequence 3 from patent US 633399. 4015 bp DNA linear PAT 29-JAN-2003

GI:28075300

Unknown.

Unclassified.

1 (bases 1 to 4015)

Monia, B. P., Gaarde, W. A. and Wanciewicz, E.

Antisense inhibition of test expression

JOURNAL Patent: US 631399-A 3 18-DEC-2001;  
ATTRES Location/Qualifiers  
source 1, 4015  
/organism="unknown"  
SE COUNT 663 a 1363 c 1275 g 714  
IGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CC CG C C C T T G A G C A C A A T G A C C C G C G C T C T C G T T G C C C C G G G T G C G C T T C T T G C T G C G 82  
40 C C C G G C C A C C C C C G G A T G C C G C G C G T C C C C G C T G C G A G C C G T G C C T C C T T G C T G C G 99  
83 C A G C C A T A C C G G A G T G T G C C C G C T G C C A A C T T T G T G C G G C G C C T G G G G C C C A G G G 142  
100 C A G C C A C T A C C G A G G T G C T C G C G T T G C G C A G T T C G T G C G G C G C T G G G G C C C A G G G 159  
143 C A G C G G C T T G T G C A A C C C G G G A C C C G A A G A T C T A C C G C A C T T T G T G T G C C C A T G C C T 202  
160 C T G C G G C T G T G C A G C G C G G G A C C C G G G G G T T T C C G C G C G T G T G G C C C A G T G C C T 219  
203 A G T G T G C A T G C A C T C G G C C T C A C A G C C T C C A C C T G C C G A C C T T T C C T T C C A C A G A G T G T C 262  
220 G G T G T G C G T G C C C T G G A G C A C G C C G C C C C C G C G G C C C C T C C T T C G C G A G T G T C 279  
263 A T C C T G A A A G A G C T G T G G C C A G G T T G T G A G A C T C T C G A G C G A A C A G A G A A A 322  
280 C T G C C T G A A G A G C T G T G G C C C A G T G C T G C A G A G C T G T G C A G C G C G C G C G A A A 339  
323 C G T G C T G G C T T T G C C T T T G A G C T G C T T A A C A G G G C C A G A G C G G C C T C C A T G G C C T T 382  
340 C G T G C T G G C C T T C G A C T T C G C C T G C T G A C G G G C C C G C G G G C C C C C C G A G C C T T 399  
383 C A C T A G T A G C G T G C T A G C T A C T T C C C C A A C A C T G T T A T T G A G A C C C T G C G T G C A G T G 442  
400 C A C C A C C A G C T G C G C A C T A C T T G C C C C A C A C G S T G A C C G A C C A C T G C G G G A G C G C 459  
443 T G C A T G E A T G C T A C T G T T G A G C G A G T G G C G A C A C C T G T G T G T C T A C C T G C T T G G C A C A 502  
460 G G C G T G G G G C T C T G C T C G C C G C G T G G C G A C G A C G T G T G T T C A C C T G C T G C G C A G 519  
503 C T G T G C T C T T A T C T T C G T G T G C C C C A C T G T G C T A C A G A G T G T G G T C T C C C C T 562  
520 C T G C G C G C T C T T T G T G T G T G G C T C C C A G C T G C G C T A C C A G T G T G C G G C G C G C T 579  
563 G T A C C A A A T T T G T G C C A C A C G A T A T C T G G C C C T C T G T C C G C T A G T T A C A G G C C C A C 622  
580 G T A C C A G C T C G C G C T G C C A C T C A G G C C C G G C C C C G C C A C A C G C T A G T G A C C C C A G 639  
623 C C G A C C C T G G C A G A A T T T C A C T A A C C T T A G G T T C T T A C A A C A G A T C A A G A G C A G T A G 682  
640 G C G T C T G G A T G C ----- G A A C G G C C C T G G A A C C A T A G 672  
683 T C G C C A G G A A G C A C G A A A C C C T G G C C T T G C C A T C T C G A G G T A C A A G A G G A C T C T A G 742  
673 C G T C A G G A G G C C G G G T C C C C T G G A C T C C A G C C C G G G T G A G A G G C C G G G G 732  
743 T C T C A C A G T A C A G T G T G C C T C A G C T A A A A G G C C A G A T C T A T C C T G T C G A G A G T 802  
733 C A G T G C C A C C C A A G T C T G C C G T T G C C A A G A G C C A G G C G T G C G C T G C C C T G A G C C 792  
803 G A G A G A G G A C C C C A C A G C A G G T G T A C C A A C C C A T C A G G C A A A T C A T G G G T G C C A G 862  
793 G A G C G A C G C C G T T G G C A G G G T C C T G G G C C A C C C G G C A G A C G C T G A C C G A G 852  
863 T C C T G C T C G G T C C C C A G A G T G C C T A C T ----- G C A G A A A G A T T G C T T C T A A 913  
853 T G A C C G T G T T C T T G T G T G T G T C A C C T G C A G A C C C C G A A G A G C C A C C T C T T T G G A 912  
914 A G A A A A G G T G T G T G A C C T G A G T C T C T C --- T G G G T C G G T G C T G T A A C A C A A G C C C A G 970

913	971	973	1028	1033	1088	1090	1148	1150	1208	1210	1268	1270	1328	1330	1343	1390	1403	1450	1463	1510	1523	1570	1583	1630	1643	1690	1703	1750	1763	1810	1823	1870	1883	1930	1943	1990	
Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy
913	971	973	1028	1033	1088	1090	1148	1150	1208	1210	1268	1270	1328	1330	1343	1390	1403	1450	1463	1510	1523	1570	1583	1630	1643	1690	1703	1750	1763	1810	1823	1870	1883	1930	1943	1990	
GGGTGGCGCTCTCTGGCAGCGCCACCTCCACCCATCCGTGGCGCCGACGACCGCGGG	CTCCACAT---CTCTGTGTCCACACCCCGCCAAAATGCTTTTTCAGCTCAGGCCATTAT	CCCCCATCCACATCGCGGCGCACAGTCCCTGGGACACGCCTTGTCCGCCCGGTGTACGC	TGAGACGAGCATTTCTTTTATCCAGGGGAGATGSCCAAGAGCGTCTTAAACCCCTCAT	CGAGACCAGGACATTTCTTACTTCTCAGGCGA---CAGAGAGAGCTGCGGCCCTCCTT	CCTACTCAGCAACCTTCAGGCTTAACCTTGACTGGGCGCAGGAGACTGTGGTGAGATCATCTT	CCTACTCAGCTCTCTGAGGCCCAGCCGTGACTGGCGCTCGGAGGCTCTGTGGAGACATCTT	TCGTGGCTCAAGGCTTAGGACATCAGGACCATCTCTGTCAGGACACACCGTCTATGCGCTCG	TCGTGGTTCAGGCCCTTGAGTCCAGGACTCCCGCAGTTGCCCGCTGCCCCAGCG	ATACTGGCAGATGGGCCCTCTGTTCCAAACAGCTGTGTTGTAACCAATGACAGATGGCAATA	CTACTGGCAAAATGGGCCCTCTGTTTCTGAGAGTGTGTGGAAACCAACGCGCATGCGCCCTA	TGTCAGACTCCTCAGGTCACTATGCAGGTTTCGAAACAGCAAAACCAACAGGTGACAGATGC	CGGGTGTCTCTCAGACGCACTGCCCGCTGGAGCTCGGTCACCCAGCAGCGGTTGT	CT-----TGAACACAGGCC	CTGTGTCCCGGAGAGCCCAAGGGCTCTGTGGCGGCCCGGAGGAGGACACAGACCC	ACCGCAGCTCATGATTTGCTCGGCTGCACAGCAGTCCCTGGCAGGTATATGTTTTTCT	CCGTGCTGTGTGAGCTGCTCCGCGACACAGCAGCCCTGCGAGGTGATACGGCTTCGT	TCGGCCTGTCTCTGAAAGTGTGTCTGTGTCTCTGGGTATCCAGGCAATATGACG	GCGGCCCTGCTGCGCGGCTGTGTGCCCGACGCTCTGGGGCTCCAGGCAACAGAACG	CGGCTCTTTAAGAACTTAAAGAGTTTCATCTCGTTGGGGAAATACGGCAAGCTATCACT	CGCTTCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCT	GCAGGACTGATGTGGAGATGAAGTAGAGGATGCCACTGGCTCCGACAGCAGCCCGG	GCAGAGCTGACTGGGAAGATGAGCTGCGGACGTGCGCTTGGCTTCGCGCAGGCGCCAGG	GAAGACCGTGTCCCGCTGCAGAGCACCGTCTGTGAGGAGAGATCCCTGGGTACGTTCT	GTTGGCTGTGTTCGGCGCAGACACATCCGTGTACGTTGTTAGTTCATTTTACATCACAGA	GTTCTGGCTGATGACACATACCTGTGTACGTTGTTAGTTCATTTTACATCACAGA	GCATGGCTGATGAGTGTGTACGTGTGTGAGCTGTCTCAGGTCTTTCTTTATGTACGGA	GAGCACATTCAGAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA	GAGCACATTCAGAGAACAGGCTCTTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA	GATTCGGCTGATGAGGCAACACCTTGGAGAGTGTGGGCTACGGGAGCTGTCCAGAGGA	ANGCATTTGGAATCAGACAGCATTTGAAGAGGTGTGAGCTGTGGGAGCTGTGGAGCAGA	GGTCAGGCATCACAGGACACCTGGCTAGCCATGTCGACAGTGTGGCTTCATCCC	CAAGCCCAACCGGCTGCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACACAGCTTT	GTCAGGCTGTGCGGGTGTGGCGGATTTGTGAACATGACTACGTCTGTGGAGCCAGACGTT	GGGCGAAGGAGCAGCCAGCATTTCAACCAAGCGTCTCAAGACTCTTCTTCAGCATGCT	CCGACAGAAAGAGGCGCCAGGCTCTCACTCGAGGGTGAAGGACACTGTTTCAGCGTCT		



```
2003 CAACTATGAGCGAGCAAAACATCTCTACCTATTAGGGTCTTCTGTACTGGGTATGAATGA 2062
2050 CAACTACGAGCGGCGCGCGCCCGGCTCTCTGGGCGCTCTGTGTGGGCTGTGACGA 2109
2063 CATCTACAGGACTGGCGGCGCTTGTGTGGGTGGGTCTCTGAGCAGACACACCGAG 2122
2110 TATCCACAGGGCTGGCGACCTTGTGTGGGTGGGTCTCTGAGCAGACACACCGAG 2169
2123 GATGTACTTTTAAAGCAGATGTGACCGGGCTTATGATGCCATCCCCAGGGTAAAGCT 2182
2170 GCTGTACTTTTCAAGTGGATGTGAGGGGCGGTACGACACCATCCCCAGGACAGGCT 2229
2183 GGTGGAGGTGTTCCTCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCCGAGTA 2242
2230 CAGGGAGGTTCATCGCCAGCATCATCA---AACCCAGAAACAGTACTGCGTGGTTCGGTA 2286
2243 TGCAGTGGTCCGAGAGATAGCAAGGCCCAAGTCCCAAGTCCCTTTAGGAGACAGGTAC 2302
2287 TGCAGTGGTCCAGAGGCCCGCCATGGGACGTCCGCAAGGCCCTTCAGAGCCACGTCTC 2346
2303 CACCTCTCTGACCTCCAGCCATACATGAGGCCAGTTCCTTAAGCATCTGCAAGATTCAGA 2362
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGGTCACTGTCAG-----GA 2400
2363 TGCAGTGGTCACTGAGGAACCTCGTGTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422
2401 GACACAGCCGCTGAGGGATGCGGTCTGTCTATCGAGCAGAGCTCTCTCTGATGAGGCCAG 2460
2423 CAGCAGCGTGTGTGACTTTCTCTGCACTTCTCTGGTCACTGCTGCTGCTGCTGCTGCTG 2482
2461 CAGTGGCGCTCTTCAGCGCTTCTCTAGCGTTCATGTGCCACACGCGGTGCCATCAGGGG 2520
2483 CAGTGGTATACAGTGGTCCAGGGCTCCCGAGGGCTCCAGCGCTATCCACCTGCTGCTG 2542
2521 CAGTCTCTAGCTGCTGAGGGATCCCGAGGGCTCCATCTCTCTGCTGCTGCTGCTGCTG 2580
2543 CAGTCTGTGTTCGAGAGCATGAGAACAGTGTGTGTGAGGTGTCAGCGGATGGGTT 2602
2581 CAGCTGTGTCTACGGCGACATGAGAACAGTGTGTGTGGGGATTCGGCGGAGCGGCT 2640
2603 GCTTTTACGTTTGTGTGATGACTTCTGTGTGTGAGCGCTCATCTGGACCAAGCAAAAC 2662
2641 GCTCCTGGGTTGTGTGATGATTTCTGTGTGTGACATCTACCTCACCCAGCGGAAAC 2700
2663 CTTCTCTCAGCACCTGCTCATGCGGCTTCTGATGATGGGTGATGATGATGATGATGATG 2722
2701 CTTCTCTCAGGACCTGCTGAGGTGCTGCTGATGATGATGATGATGATGATGATGATG 2760
2723 GACAGTGGTGAATCTCTCTGTGAGCTGCTGATCCCTGGGTGTGAGCTCCATACAGCT 2782
2761 GACAGTGGTGAATCTCTCTGTGAGACGAGGCGCTGGGTGTGAGCTCTTGTGATGAT 2820
2783 GCTGTCTCAGTCTGCTTCTCTGTGTGTGCTGCTGTGTGAGCTCTGAGCTTTGGAGGT 2842
2821 GCGGCGCCACGCGCTATTCTCTGTGTGTGCTGCTGTGTGATACCGGACCTTGGAGGT 2880
2843 GTTCTGTGATCTCTCAGGTTATGCCCAGACCTCAATTAAGACGAGCTCACTTCCAGAG 2902
2881 GCAGAGCGACTACTCTCAGCTATGCCCGGACCTCCATCAGAGCGAGTCTCACTTCAACCG 2940
2903 TGTCTTCAAGCTGGGAGACCATGCGGACAGCTCTCTGTGTGTGTGCTGCTGCTGAGTG 2962
2941 CGGCTTCAAGCTGGGAGGAAATGCTGTGCAAACTCTTTGGGGTCTTTCGGCTGAAGTG 3000
2963 TCAGGCTTATTTCTAGACTTGCAGGTGACACAGCTTCCAGACAGTCTGCATCAATATATA 3022
3001 TCAGGCTTGTTCGATTTGAGGTGACAGCTTCCAGAGCGGTGTGCAACATCTA 3060
3023 CAGATCTTCTGCTTCCAGGCTTACAGGTTCCATGATGTGTGATTCAGCTTCCCTTTGA 3082
3061 CAAGATCTCTCTGTCAGGCGTACAGGTTTTCAGGATGTGTGTGTGCTGCTGCTGCTTCA 3120

3083 CAGCGCTTTAGGAGAACCTCACAATCTTTTGGGATCATCTCCAGCCAGCATCTG 3142
3121 TCAGCAGATTTTGGAGAACCCCAATTTTCTGCGGTGTCATCTGTGACACGCGCTCCCT 3180
3143 CTGCTATCTCTATCTCTGAGGTCAAGATCCAGGAATGACACTAAAG-----GCCTC 3193
3181 CTGCTACTCTCATCTCTGAAAGCCAGAACAGGAGTGTCTGCTGGGGCCAAAGCGCGC 3240
3194 TGGCTCTCTTCTCTCTGAGCGGCACATTTGGCTCTGTACAGGCTTCTGCTCAAGCT 3253
3241 CCGCCCTCTGCGCTCCGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
3254 GGTCTCTATCTCTGCTCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313
3301 GACTCGACACCGTGTCACTTACGTGCGACCTCTCTGGGTCACTCAGGACAGCCAGAGCA 3360
3314 GCTGTGCCGAGCTTCCAGAGGCGCAATGACCATCTTAAAGCTGCACTGACCCAGC 3373
3361 GCTGAGTCGGAAGCTCCCGGGAGCGAGCTGACTGCTGAGGCGCGCAGCAACCCGCG 3420
3374 CCTAAGCAGACACTTTCAGACCATTTTGGACTAA 3407
3421 ACTGCTCTGAGACTTCAAGACCATCTCTGAGCTGA 3454

RESULT 13
LOCUS AR265996 4015 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6492171.
ACCESSION AR265996
VERSION AR265996.1 GI:29694842
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4015)
AUTHORS Monia,B.P., Gaarde,W.A., Preter,S.M. and Wanciewicz,E.
TITLE Antisense modulation of TERT expression
JOURNAL Patent: US 6492171-A 3 10-DEC-2002;
FEATURES
Location/Qualifiers
1..4015
source
BASE COUNT 663 a 1363 c 1275 g 714 t
ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;
Best Local Similarity 68.9%; Pred. No. 0;
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGGGCTTTGAGCACAATGACCCGGCTCTCTGTTGCTGCGCGCTGCGCTGCGCTGCGCTCTGCTGCG 82
DB 40 CCGGGCCACCCCGCGATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTGCGCTGCGCTGCGCTGCG 99
QY 83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCCGAGGG 142
DB 100 CAGCCACTACCGAGGTGTGCTGCGCTGGCCACCTGCTGCGGCGCTGGGCGCCCGAGGG 159
QY 143 CAGGCGCTTGTGGAACCCCGGGACCCGAGATCTACGCACTTTGTTGTTGCTGCTGCTGCTGCT 202
DB 160 CTGGCGCTTGTGTCAGCGCGGGACCCCGCGGCTTTCGCGCGCTGCTGCGCGCTGCTGCTGCT 219
QY 203 AGTGTGATGCACTGGGCTCAGAGCTTCCAGCTTCCAGCTTCTTCCAGCAGGTGTC 262
DB 220 GGTGTGCTGCTGCTGAGACGACGCGCGCGCCCGCGCGCTTCTTCCGCGCAGGTGTC 279
QY 263 ATCCCTGAAGAGCTGTGGCGAGGTTGTGAGAGACTCTGCAAGACTCTGCGAGCGCAACGAGAGAA 322
DB 280 CTGCTGAAGAGCTGTGGGCGCGAGTGTCTGAGAGGCTGTGCGAGCGCGCGCGCGAGAA 339
QY 323 CGTGCTGCTTGTGCTTGTGCTGCTTAAAGAGCGCGAGGCGGCTTCCCATGCGCTT 382
DB 340 CGTGCTGCGCTTTCGCGCTTCTGCGCTGCTGAGCGGGCGCGCGCGCGCGCGCGCTT 399
```



383 CACTAGTAGCGTGGTGTAGCTACTTGGCCCAACACTGTATTAGAACCTCGTGTCAAGTGG 442  
400 CACCACAGCGTGGCGAGCTACTTGGCCCAACACGGTGAACGACACTCGGGGAGCGG 459  
443 TGCATGATGCTACTGTGTAGCGGAGTGGCGAGCACTGCTGTCTACTCTGCTGGCACA 502  
460 GCGGTGGGGGTGCTGTCTGGCGCGCGTGGCGAGCAGCTGTGTTTCACTGTGTGSCAG 519  
503 CTGTGCTCTTTATCTTCTGTGTGCCCCCAGCTGTGCTACAGGTGTGTGGTCTCCCT 562  
520 CTGCGCGCTCTTTGTGTGTGTCTCCAGCTGCGCTACCAAGTGTGCGGCGCGCT 579  
563 GTACCAAAATTTGTGCAACAGGATATCTGGCGCTCTGTGTCTGCTGTAGTGTACAGGCCAC 622  
580 GTACAGCTCGCGCTGCCACTCAGGCGCGCGCGCGCGCGCAACGCTAGTGGACCCCGAAG 639  
623 CCGACCCGTGGGCGAGGAATTTCACTAACCTTAGGTTCTTAAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GACGGGCGCTGGNACCATAG 672  
683 TCGCCAGGAAGACACGAAACCCCTGGCGCTTGCATCTCGAGGTACAAAGGCACTGTAG 742  
673 GGTCAAGGAGGCGCGGCTGCCCTTGGCGCTGCCAGCCCGGCTGCGAGGAGGCGCGGG 732  
743 TCTCACAGTACAGTGTGCTTCCAGTAAAGGCGCAGATGCTATCTCTGCCGAGGT 802  
733 CAGTCCAGCGGAAGTCTGCGGTTTGCACAGAGGCCAGCGCTGGCGCTGCCCTGAGCC 792  
803 GGAGGAGGACCCACAGGAGGTGCTACCAACCCCATCAGGCAATCATGGGTGCCAAG 862  
793 GGAGCGGACGCGGTTGGCGAGGCTCTTGGGCGCACCGCGGAGGACGCGTGGACCGAG 852  
863 TCCTGCTCGGTCCCGGAGGCGCTACT-----GCAGAGGAATTTGCTCTTAA 913  
853 TGACCGTGTCTGT 912  
914 AGGAAGGTGTCTGACCTGAGTCTCTC---TGGGTGCGTGTGTGTGTGTGTGTGTGTGT 970  
913 GGTGCGCTCTCTGGCAGCGCCACTCCACCCATCGTGGCGCGCGCAGCACACGCGGG 972  
971 CTCACAT---CTGT 1027  
973 CCCCCCATCCCATGTGCGGCGCACCATGCTTGTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACCAAGATTTCTTACTCCAGGGAGATGGCAAGAGCGCTTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1089  
1088 CCTACTCAGCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1147  
1090 CCTACTCAGCTCTCTTGGGCGCGCGCTGACTGGCGCTCGGAGGCTCTGTGAGACCATCT 1149  
1148 TCTGGGCTCAAGGCGCTAGGACATCAGGACCACTCTGAGGACACACCGCTCTATCGGCTG 1207  
1150 TCTGGGTTCCAGGCGCTGT 1209  
1208 ATACTGGCAGATGGCGCGCTTGTTCACAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1267  
1210 CTACTGGCAAAATGGCGCGCTTGTTCGAGCTGCTTGGGAAACCAACGCGCAGTGGCGCTTA 1269  
1268 TGTACAGTCTCTCAGTGCATTCGAGGTTTGAACAGCAAAACCAAGGTGTACAGTGC 1327  
1270 CCGGTGTCTCTCAAGCGCACTGCGCGCTGCGAGCTGCGGTACCCCGACGACCGCGGTGT 1329  
1328 CT-----TGAAACACCGAGCC 1342  
1330 CTGTGCGCGGAGAGCCCGAGGCTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1389  
1343 ACCGACCTCATGATTTGTCTCCGCTGCAAGCAGTCCCTTGGCAGGTATATGGTTTCT 1402  
1390 CCGTGGCTGTGTGCACTGCTCCCGCAGCAAGCAGCGCGCTTGGCAGGTGTACGGCTGT 1449  
1403 TGGGCGCTCTCTGCAAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1462

1450 GCGGGCTGCTGCTGGCGCGCTGTGTCGCCCAAGGCTCTTGGGCTCCAGGCAACGAAAG 1509  
1463 CCGCTTCTTTAAGAACTTAAAGAAATGCTATCTGTTGGGAAATACGCAAGCTATCTACT 1522  
1510 CCGTTCCTCAGGAACACCAAGAAATGTTCTTCTTGGGAGCATGTCRAAGCTCTCGCT 1569  
1523 GCAGGAATGATGTGGAAGATGAAAGTAGAGATTTGCCACTGGCTCCGACAGCCCGGG 1582  
1570 GCAGGAGCTGACGCTGGAAGATGAGGCTGCGGAGCTGCGCTTGGCTGCGCAGAGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGACAGACACGCTCTGAGGAGAGGATCTCTGGTGTAGCTTCT 1642  
1630 GGTGTGTGTGTTCGCGCGCAGAGACCGCTCTGCTGAGGAGATCTCTGGCAGATTTCT 1689  
1643 GTTCTGGCTGTAGGACACATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1702  
1690 GCACTGGCTGTAGT 1749  
1703 GAGCACAATTCAGAGAAACAGGCTCTTCTTCTACCTGTAAGATGTGTGAGGAGCTGCA 1762  
1750 GACCACTGTTTCAAGAAACAGGCTCTTCTTCTACCGAAGATGTGTGAGCAAGTGTCA 1809  
1763 GAGCATTTGAGTGTGAGCAACACCTTGAGAGATGCGGCTTACCGGAGCTGTCAACAGAG 1822  
1810 AAGCATTTGATCAGACAGCACTTGTGAGAGGCTGTGAGCTGCGGAGCTGTGGAAGCAG 1869  
1823 GGTGAGCATCACAGGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1882  
1870 GGTGAGGAGCATCGGGAAGCAGGCGCGCTCTGTGAGCTGTGAGCTGTGAGCTGTGTGT 1929  
1883 CAAGCCCAAGCGCTCGCGCCCATTTGTGAACATGATGATGTATAGCATGGGTGTACAGAGCT 1942  
1930 CAAGCTGTAGCGGCTGCGGCGCATTTGTGAACATGAGCTGTGAGCTGTGAGGAGCAG 1989  
1943 GGGCAGAGGAAGAGCGCGCGCATTTGTGACCGGCTGTCAAGATCTCTTCTTGTGTGTGT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGCGCTCTACCTCGAGGCTGTGAGGCACTGTTCAGCGTGT 2049  
2003 CAATATGAGCGGACAAAACATCTCACCTTATGGGTCTTCTGTACTGGGTGTATGAATGA 2062  
2050 CAATCTAGCGCGGCGGCGCGCGCTCTGTGGCGCTCTGTGTGTGTGTGTGTGTGTGTGT 2109  
2063 CATCTAGAGCTGTGCGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
2110 TATCCAAGGCGCTGTGCGCACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
2123 GATGTACTTTTGTAAAGCGAGTGTGACCGGCGCTTATGATGCCATCCCGCAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGTGTGACGGCGCGCTGTACGACCATCCCGCAGGACGCT 2229  
2183 GGTGAGGTTGTTCCAATATGATCAGGCACTCGGAGAGCAGTGTGTATCCGCGATGA 2242  
2230 CACGAGGTTCTATCGCCAGCATCATCA---AACCACCAACACGTACTGCTGCTCGGTA 2286  
2243 TGCATGCTCGGAGAGATAGCCAGGCAAGTGTCAAGTCTTCTTGTAGGAGACGCTCATC 2302  
2287 TGGCGTGTTCAGAGGCGCGCGCATGCGGCACTGTGCGGCGCGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACTCTCAGCCATACATGCGGCGAGTTCCTTAAGCATCTGTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGAGATTCGTGGCTCACCTGTGAG---GA 2400  
2363 TGCAGTGTCACTGAGAACTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2422  
2401 GACCAGCGCGCTGAGGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460  
2423 CAGCAGCTGT 2482  
2461 CAGTGGCGCTTCTGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2520  
2483 CAGGTGTATACGAGTGTGCGAGGCGATCCCGCAGGCTTCAAGCTTATCAGCTGTGTGTGT 2542

2521 CAAAGTCTACGTCAGTCCAGGGATCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCGTGTTTCGGAGACATGAGAGCAAGCTGTTCGTAGGTGCGAGCGGATGGTT 2602  
2581 CAGCCTGTGCTACGGCGACATGAGAACAGCTGTTCGGGGATTCGGCGGGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTGTGTGACGCTCCTCATTGGACCAAGCAAAAC 2662  
2641 GCTCCTGGTGTGTGATGATTTCTTGTGTGACACCTCCTCCTCCACCGCGGAAAC 2700  
2663 CTTTCTCAGACCCCTGTCATGCGGTTCTCAGTATGGGTGATGATAAACTTCAGAA 2722  
2701 CTTTCTCAGACCCCTGTCGAGGTTCCTCAGTATGGGTGATGATGCGGTGCTGCGAA 2760  
2723 GACAGTGGTGAACCTTCCCTGTGAGCTCGTACCTCGGTGGTGGTGCAGTCCATACCACT 2782  
2761 GACAGTGGTGAACCTTCCCTGTAGACAGGCGCTGGGTGGCGCTTTTGTTCAGAT 2820  
2783 GCTGTCTACGCTGCTTTCCTGTGGTGTGCTGTGTGACATCAGACTTTGGAGGT 2842  
2821 GCGGGCCACGGCTATTCCCTGTGGTGGCTGTGTGTGATACCGGACCTCGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTATGCCCCAGACCTCAATTAAGACGAGCTCCTTCCAGAG 2902  
2881 GCAGAGCACTACTCAGTATGCCCGACCTCCATCAGAGCACTCCTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAAGACCATGCGAACAGCTCCTGTGGTCTTTGCGTTGAAGT 2962  
2941 CGGCTTCAAGGCTGGGAGGAAATGCGTGGGCTTCTTGGGGTCTTGGCGTGAAGT 3000  
2963 TCACGCTATTTCTAGACTTCAGGTGACAGCTCCAGAGCTCCAGAGCTTCGATCAATATA 3022  
3001 TCACGCTGTTTCTGATTTGAGTGTGAGTGAACAGCTCCAGAGCTGTCACCAATCTA 3060  
3023 CAAGATCTTCTCTGCTCAGGCTTACAGGTTCATGATGATGATGATGATGATGATGAT 3082  
3061 CAAGATCTTCTCTGCTCAGGCTTACAGGTTCATGATGATGATGATGATGATGATGAT 3120  
3083 CAGAGTGTAGAGAAAGCTTACATTTCTTGGGATCATCTCCAGCCAGATCTCTG 3142  
3121 TCAGCAAGTTTGAAGAAACCCACATTTTCTGCGCTCATCTCTGACAGGCTCTCT 3180  
3143 CTGCTATGCTATCTCAAGGTCAAGATCCAGGATGACATCAAG-----GCTC 3193  
3181 CTGCTACTCATCTCTGAGCCAGAACCCAGGATGTCGTGGGGCCAGGGCGCGC 3240  
3194 TGCTCTTCTCTGAGCCGACATTTGCTCTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTCTGCTGAGCCGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGCTGCTCATCTGCTATACAAATGCTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313  
3301 GATCGACACCGTGTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCGGAGCTCCAGAGCGACATGACCATCTTAAAGTCTGAGTGAACCCAGC 3373  
3361 GCTGAGTGGAGCTCCCGGACAGCTGCTGCTGAGGCGGAGGCGGAGCAACCCGCG 3420  
3374 CTTAAGCAGACACTTTCAGACCACTTTGCACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGAGCTGA 3454

ULT 14  
19310  
US  
INITION  
AX019310  
Sequence 10 from Patent WO9940221.  
AX019310  
AX019310.1  
GI:10043297  
WORDS  
RCE  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 4015  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCGGCTTGAGCACAATGACCCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGTGCG 82  
DB 40 CCGGCAACCCCGCGATGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGTGCG 99  
QY 83 CAGCCGATACCGGAGGTGCGCGCTGGCAACCTTTGCGGCGCTGGGCGCCGAGG 142  
DB 100 CAGCCACTACCGGAGGTGCTGCGCTGGCCACGTTCTGCGGCGCTGGGCGCCAGG 159  
QY 143 CAGGCGCTTGTGCAACCCCGGAGCCCGAAGATCTACCGCACTTTGTTGCCCAATGCT 202  
DB 160 CTGGCGCTGTTGTCAGCGCGGAGCCCGCGGCTTCCGCGCGCTGTTGCCCAATGCT 219  
QY 203 AGTGTGATGCTCTGGGCTCAGAGCTTCTGCTGCGAGCTTTCTTCCACAGGTGTC 262  
DB 220 GGTGTGCGTGGCTTGGGACGACCGCGCGCGCGCGCGCGCTTCTTCCGCGAGGTGTC 279  
QY 263 ATCCCTGAAAGAGCTGTTGGCGAGGTTGTGAGAGACTCTGCGAGCGCAACAGAGAAA 322  
DB 280 CTGCTGAGGAGCTGTTGGCGCGAGTGTGCGAGCGCTGTCAGCGCGCGCGAGAA 339  
QY 323 CTTGTGCTTGTGCTTGTGAGTGTGTTAAGAGCGCAGAGCGCGCTTCCATGGCTT 382  
DB 340 CTTGTGCTTGTGCTTGTGAGTGTGTTAAGAGCGCGCGCGCGCGCGCTTCCATGGCTT 399  
QY 383 CACTAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442  
DB 400 CACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
QY 443 TGCATGATGCTTACTTGTGAGCGAGTGGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 502  
DB 460 GCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
QY 503 CTGCTCTTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562  
DB 520 CTGCGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
QY 563 GTACCAAAATTTGTGCGACCGAGTATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622  
DB 580 GTACCAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639  
QY 623 CCGACCGGTGGCGAGGAAATTTCACTAACCTTAGTGTTCACACAGATCAAGAGCAGTAG 682  
DB 640 GCGTCTGGGATG-----GAACGGGCTTGGAAACCATAG 672  
QY 683 TCGCAGGAGACACCGAAACCCCTGGCTTCCATCTCGAGGTACAAAGAGGATCTGAG 742  
DB 673 CTTAGGAGGCGCGGCTTCCCTTGGGCTGCTGCGAGCCCGGCTGCGAGGAGCGCGGGG 732  
QY 743 TCTCAACAGTACAAGTGTGCTTCAAGTGAAGAGCGAGATGCTATCTCTGTCGAGAGT 802  
DB 733 CAGTCCAGCGGAGTCTGCGCTTGGCCAGAGGCGCGAGGCGTGGCGCTGCGCTGAGCC 792  
QY 803 GGAGGAGGAGCCACAGGAGGTGCTTACCAACCCCATCAGGCAATCATGCTGCTGCTGCTGCT 862

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Dahn, M.W., Phelps, R.C. and Brockmeyer, C.  
Method for quantitatively analyzing tumor cells in a body fluid and  
test kits suited therefor  
Patent: WO 9940221-A 10 12-AUG-1999;

DAHM MICHAEL W (DE); PHELPS ROBERT C (DE); BROCKMEYER CARSTEN (DE)  
Location/Qualifiers

793 GGAGCGAGCGCGCTTGGCGAGGGCTCTGGGCCCAACCGCGGCAGAGACGCGTGGACCGAG 852  
863 TCCTGCTCGGTCCCGCGAGGTGCCTACT-----GCAGAGAAAGATTGTCTCTTAA 913  
853 TGACCTGGTTCTGTGTGTGTACCTGCGAGACCGCGGAGAGACCACTCTTTGGA 912  
914 AGAAAGGTGTCTGACCTGAGTCTCTC-----TGGGTGGTGTGTCTGTAAACAAGACCGAG 970  
913 GGGTGGCTCTCTGGCACGCGCACTCTCCCAACCATCCGTGGGCCCGCCAGCACACCGGGG 972  
971 CTCACAT---CTCTGCTGTACCAACCGCGCAAAATGCTTTACAGCTCAGGCCATTTAT 1027  
973 CCCCCCATCAATCGCGGCACCAAGTCTCTGGACACGCTTGTCTCCCGGTGTAGCG 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAAGAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCAGCGA---CAAGAGCAGCTGCGCGCTCTCT 1089  
1088 CTTACTCAGCAACTCCAGCTTAACTTTGACTGGGCCAGGAGACTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCAGGCCCGAGCTGACTGGCGTCGGAGGCTCGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCCCTAGACATCAGAACCACTCTGCAGGACACACCGTCTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCCCTGGATGCCAGGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCG 1209  
1208 ATACTGGCAGATCGCGCCCTGTTTCCACAGCTGTGTGTGAACCATGACAGTGTGCAATA 1267  
1210 CTACTGGCAATCGCGCCCTGTTTCTGGAGCTGTGTGGAAACCAACGCGCAGTGCCTTA 1269  
1268 TGTCAAGACTCTCAGGTCAATTCAGAGTTTTCGAAACAGCAAAACCAAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACGACTGCCCGCTGGAGCTGGGTTCACCCCGCAGCGCGTGT 1329  
1328 CT-----TGAACACAGGCC 1342  
1330 CTGTGCGCGGAGAACCCCGAGGCTCTGTGCGGCCCGCCGAGGAGGAGACACAGACCC 1389  
1343 ACCGCACTCATGATTTGCTCGCGCTGCACAGCAGTCCCTCGCAGGTATATGTTTCT 1402  
1390 CGTGGCTGGTGCAGTCTCCGCGCAGCAGACAGACCCCTCGCAGGTGTACGGCTTGT 1449  
1403 TCGGGCTGTCTCTCAAGTGTGTCTGTCTGTCTGTGGGTACCAAGGCAATGAGCG 1462  
1450 GCGGCTGCTGCGCGGCTGTGTGCGCCCGAGGCTCTGGGCTCCAGGCAACAGAACG 1509  
1463 CCGCTCTTTAAGAACTTAAAGATTCTCTGTGGGAAATACGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAGCTCTGCT 1569  
1523 GCAGGAAGTATGTGAAGATGAAGTAGAGGATTGCCACTGGCTCCGACAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGACCTGGCTTGGCTGCCAGAGGCCAAG 1629  
1583 GAAGGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCTCGGTACGTTCT 1642  
1630 GGTGGCTGTGTTCGGCGCGCAGAGCACCGTCTGGGTGAGGAGATCTGGGCCAAGTCTCT 1689  
1643 GTTCTGGCTCATGGACACATACGTGTGTACAGTGTGTAGGTCAATCTTTTATACATCAGA 1702  
1690 GCACCTGGCTGATGAGTGTGTACGTCTGTGAGTGTGTCTCAGTCTCTTTTATGACGGA 1749  
1703 GAGCAATTCAGAGAAAGAGGTCTTCTTACCGTAAAGTGTGTGGAGGACAGCTGCA 1762  
1750 GACCAGTTTCAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGATTGGAGTCAGGCAACACCTTCAGAGAGTGGCGCTACGGGAGCTGTACAGAGGA 1822  
1810 AAGATTGGATTCAGACAGCACTTGAAGGGTGTGAGCTGCGGAGCTGTGGAGCAGA 1869  
1823 GGTCAAGCATCAACAGACACTGGGTGAGCAATGCCCATCTGCAGACTGCGGTTCATCCC 1882  
1870 GGTCAAGGCTCGGAGAGCAGGCGCGCTGTGTGAGCTCCAGACTCCGCTTCATCCC 1929

1883 CAAGCCCAACCGCTCGCGGCCCAATGTGAACATGAGTTATAGCATGGTACCAGAGCTTT 1942  
1930 CAAGCCTGACGGGTGCGGCGGATTTGTGAATGAGACTACGTCTGTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGAGGAGCAGCGCCAGCAATTTCAACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CGGCAGAGAAAGAGGCGCGAGCGTCTCACTCTGAGGGTGAAGGCACTGTTTCAAGCTGCT 2049  
2003 CAACTATGAGCGGCAAAACATCTCTCACCTTATGGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACTACGAGCGGCGCGCGCGCCCGGCTCTCTGGGCGCTCTGTGTGCTGGCGCTGAGCGA 2109  
2063 CATCTAAGAGACTGGCGGCGCTTGTGTCTGGCTGTGGTGTGGTCTGACCCAGACACCCAG 2122  
2110 TATCCACAGGCGCTGGCGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
2123 GATGTACTTTTGAAGGAGATGTGACCGGCGCTTATGATGCCATCCCGCCAGGGTAACT 2182  
2170 GCTGTACTTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTTGTGCAATATGATGATGAGCACTCGGAGAGCAGTACTGTATTCGCCAGTA 2242  
2230 CACGAGGTCATCGCAGCATCATCA---AACCCAGACACACGTACTGTGTGTGGTACGTA 2286  
2243 TGCAGTGTTCGGGAGATAGCCAAAGGCCAAGTCCCAAGTCTCTTTAGGAGACAGGTCAC 2302  
2287 TCCGTGTTCAGAGGCGCGCCCATGGGACGTCGCCAAGGCTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTCTCTTAAAGCATGTCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTATATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGCATGAGGACTCCGTTGATCTGACGACAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCGACCGCTGAGGAGTGCCTGCTCATCGACAGAGTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCCTGTTTGACTCTTCTCTGCACTCTCTCTCGCTCACAGTGTCTTAAAGATTGGTGA 2482  
2461 CAGTGGCTCTTTCGAGGCTCTTCTACGCTTCATGTGCACACGCGCTGGGCTACAGGG 2520  
2483 CAGTGTCTATGACAGTGCAGGCGCATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCACGTCAGTGCAGGGGATCCCGCAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGGAGACATGAGAACAAAGCTGTTTGTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCCTGTGCTACGCGACATGAGAACAAAGCTGTTTTCGCGGATTCGCGGAGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGATCTTGTGTGTGACGCTTCTGTGAGTGTGGTGTGATGATAAATTCAGAA 2662  
2641 GCTCTGCTGTTTGTGATGATCTTCTTGTGAGTGTGATGATGATGATGATGATGATGATGAT 2700  
2663 CTTCTCTCAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
2701 CTTCTCTCAGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2723 GACAGTGTGAACTTCTCTGTGAGGCTGTGTATCCCTGGGTGGTGCAGCTCCATACAGCT 2782  
2761 GACAGTGTGAACTTCTCTGTGAGGACGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCACGCGCTTATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACTCTCAATTAAGACGAGCTCACCTTTCAGAG 2902  
2881 GCAGAGGCACTACTCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCAGCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGAGAGACCATCGGAAACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CCGCTTCAAGGCTGGGAGGAAACATGCTGCTGCGAAACTCTTTTGGGCTCTTGGCTGAACTG 3000

2963	TCACGGTCTATTCTTAGACTTGAGGTGAACAGCCTCCAGACGATCTGCATCAATATA	3021
3001	TCACAGCCTCTTTCTGGATTGCGAGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA	3060
3023	CAAGATCTTCCTGCTTCAGGCCCTACAGGTTCCATGCATGTGTGATTGAGCTTCCTTTTGA	3082
3061	CAAGATCCTCTGCTGCAGGGCTACAGGTTTCAAGCATGTGTGCTGCACTCCCATTCCA	3120
3083	CCAGCGTGTGTAGAAGAACTTCACTTTCTTGGGCAATCATCTCCAGCCAAAGCATCTGT	3142
3121	TCAGCAAAGTTTGGAAAGAACCCCAATTTTTCTGTGGCGTCACTCTGTGACGGCTCCCT	3180
3143	CTGTATGCTATCTCAAGGTCAAGAAATCCAGGAATGACACTTAAG-----GCTC	3193
3181	CTGCTACTCCATCTGANAGCCAGAACCGAGGATGTGCTGGGGCCAGAGGGCGCGC	3240
3194	TGGTCTCTTTCTCTGAAGCGCAATTTGGTCTGTCTWACAAGGCTTTCTGTCTCAAGCT	3253
3241	CGGCCCTCTGCTCCGAGGCGGTGAGTGTGGTGTGCCAACAAAGCATTTCTGTCTCAAGCT	3300
3254	GGTGTCTATTCTGTGTCATCTACAATATGCTCTGTGGAGCTCTGAGGACAGGCCAAAACCT	3313
3301	GACTGCACACGGTGTCACTTACGTGCCATCTCTGGGTCACTCAGGACAGCCAGACGCA	3360
3314	GCTGTGCGGAAAGCTCCAGAGCGCAATAGCAATCTTTAAAGCTGCAGCTGACCCAGC	3373
3361	GCTGAGTCGGAAGCTCCGGGACGACGCTGACTGCTCCCTGGAGGCGCGAGCAACCCGGC	3420
3374	CCTAAGCACAGACTTTTCAGACCATTTTTGACTAA	3407
3421	ACTGCCCTCAGATCTTCAGACCATCTGACTGA	3454
<hr/>		
JLT 15		
IS	AXI33979	4015 bp DNA linear PAT 15-MAY-2001
SSION	Sequence 1 from Patent EP1108789.	
ION	AXI33979	
IONS	AXI33979.1 GI:14139919	
CE	Homo sapiens (human)	
IGANISM	Homo sapiens	
RENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
THORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	
FILE	Chang,S.Y. and Santini,C.D.	
JURNAL	Quantitation of htert mrna expression Patent: EP 1108789-A 1 (20-JUN-2001; F. HOFFMANN-LA ROCHE AG (CH))	
URES	Location/Qualifiers	
source	1..4015 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	714 t
I COUNT	663 a 1363 c 1275 g	
FIN		
ery Match	45.8%; Score 1600.8; DB 6; Length 4015;	
st Local Similarity	58.9%; Pred. No. 0;	
atches 2379; Conservative	0; Mismatches 967; Indels 108; Gaps 9;	
23	CCCGCGCTTGAGCAACAATGACCGCGCTCTCGTTGCCCGCGGTGCGCTCTGTCTGCGG	82
40	CCCGGCAACCCCGCGATGCGCGCGCTCCCGCTGCCGAGCGTGTGCGCTCCCTGTCTGCG	99
83	CAGCCGATACCGGAGGTGGCGCTGCGCAACCTTTGTGGCGCGCTGGGGCCCGAGGG	142
100	CAGCACTACCGGAGGTGTGCGCTGCGCACGTTGTGCGCGCTTGGGGCCCGAGGG	159
143	CAGCGGCTTGTGAACCCCGGAGCCGGAATCTACCGCACTTTGTTGCCCAATGCT	202
160	CTGCGGCTGTGTCAGCGGGGACCCGCGCTTTTCGCGCGCTGTGTGGCGAGTGCCT	219

Qy	203	AGTGTGCATGCACTGGGGCTCA	CAGAGCTCCACTGCCGACCTTTCTCTTCCACCGAGTGTCTC	262
Db	220	GGTGTGCTGTCCTTGGACGCA	CGGCGCGCCCGCCGCTCTTTCGCCACAGTGTCTC	279
Qy	263	ATCCCTCGAAGAGAGCTGGTGG	CGCAGGCTGTGCAGAGACTCTGCAGCGCAACGAGAGAAA	322
Db	280	CTGCTGAAGAGAGCTGGTGG	CGCGAGTGTGCAGAGGCTGTGCAGCGCGCGCGCGAAGAA	339
Qy	323	CGTGTGCTTTTGGCTTTTGG	CTGCTTAAACGAGGCCAGAGCGCGGCTCTCCATTTGGCCTT	382
Db	340	CGTGTGCTTTTGGCTTTTGG	CTGCTTAAACGAGGCCAGAGCGCGGCGCGCGCGCGGCGCTT	399
Qy	383	CACTAGTAGCGTGGTACTCT	TGCGCAACTGTATTGAGACCTTCGCTGTGAGTGG	442
Db	400	CACCAACAGCGTGGCAGCTAC	CTGCGCAACGCTGACCGACGCACTGCGGGGAGCGG	459
Qy	443	TGCAATGATGCTACTGTTG	AGCCGAGTGGGCGACACCTGTGTGCTTCACTGCTGGCACA	502
Db	460	GSCGTGGGGCTGCTGCTG	CGCGCGTGGCGACGACGTGCTTCACTGCTGGCAGC	519
Qy	503	CTGTGCTCTTTATCTTCTG	TGTGCGCCCGCAGCTGTGCTTACGAGTGTGTGGTCTCCCT	562
Db	520	CTGCGCGCTCTTTGTGTG	TGCTCCAGCTTCCAGTGTGCGCGCGCGCT	579
Qy	563	GTACCAAAATTTGCCAC	CACGATATCTGGCGCTCTGTGTCGCTAGTTACAGGCCAC	622
Db	580	GTACAGCTCGCGCTGCACT	CAGGCCCGGCGCCCGCACACGCTAGTGGACCCGAG	639
Qy	623	CGGACCCGTGGCAGGAATTT	CACTAACCTTAGGTTCTTAAACAGATCAAGACAGTAG	682
Db	640	GGCTGTGGGATGC-----	-----GAAACGGGCTCGAAACCATAG	672
Qy	683	TGCGCAGGAACGACCGAA	ACCCCTGGCTTCGATCTCGAGTACAAGAGGCATCTGAG	742
Db	673	CGTCAGGAGCGCGGCTCC	CTGGCTGCGAGCGCGGTGCGAGGAGCGCGGCGG	732
Qy	743	TCTACCAAGTCAAGTGTGCT	TTCAGCTAAGAGGCCAGATGCTATCTCTGTCCGAGAGT	802
Db	733	CAGTGCCACGCGAAGTCTG	CCGCTGTCGCAAGAGGCCAGGCGTGCCTCTGAGCC	792
Qy	803	GGAGGAGGACCCACAGG	CAGGTGCTACCAACCCCATCGCAAAATCATGGGTGCCAAG	862
Db	793	GGAGCGGACGCCCTTGG	CAGGGGTCTGCGGCCACCCCGGCGAGACGCGTGGACCGAG	852
Qy	863	TCCTCTCGGTCCCGAGTG	CGCTACT-----GCAGAGAAAGATTTGTCTTCTAA	913
Db	853	TGACCGTGTCTTCTGTGTG	TGCTACCTGCCAGACCCCGGAAGGCCACTCTTTTGA	912
Qy	914	AGGAAAGTGTGACCTGAGT	CTCTC-----TGGGTGCGTGTGCTGTAAACAAGACCCAG	970
Db	913	GGTGTGCTCTCTTGGCAG	CGCACCTCCACCTTCGCGCGCGCAGCACCGCGGG	972
Qy	971	CTCCACAT---CTGTGTGT	CACACCCCGCAAAATGCTTTTTCAGCTTCAGGCGCATTTAT	1027
Db	973	CCCCCATCCACATCGCGG	CCACCACTCTGGGACAGCGCTTGTCCCGGTGTACGC	1032
Qy	1028	TGAGACACAGATTTTCTT	CTTACTTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT	1087
Db	1033	CGAGACCAAGCACTTCTCT	TACTCTCTCAGCGCA---CAAGAGCAGCTGCGCGCGCTCCTT	1089
Qy	1088	CTACTCAGCAACCTCCAG	CCTAACTTGCTGGGCCAGGAGACTGTGTGGAGATCACTTT	1147
Db	1090	CTACTCAGCTCTCTGAGG	CCACGCTGACTGGCTCGGAGCTCTGTGTGGAGACCATCTT	1149
Qy	1148	TCTGGGCTCAAGGCTTAG	GACATCAGGACCACTCTGCGAGGACACCGTCTATCGCGTCG	1207
Db	1150	TCTGGGTTCCAGGCGCTG	GATGCGGAGCTCCCGCAGGTTGCCCGCGCTGCCCGACGC	1209
Qy	1208	ATACTGGAGAGTGGCGCC	CTGTTCCAAACAGGTGTGTGTAACCATGAGAGTGCATA	1267
Db	1210	CTACTGGCAATGGCGCC	CTGTTCTGGAGCTGTGTGGGAACACACGCGAGTGCCTTA	1269
Qy	1268	TGTCAGATCTCCTCAGGT	CTCATTTGCAAGTGTTCGAAACGCAACCAACAGGTGTGACAGATGC	1327

1270 CGGGGTCTCTCAAGACGACTGCCCGCTGCGAGTGGGTCAACCCAGCAGCGCGTGT 1329  
1328 CT-----TGACACACGCC 1342  
1330 CTGTGCCGGGAGAACCCAGGCTCTGTGGGGCCCCCGAGGAGGAGACAGACCC 1389  
1343 ACCGCACCTCATGGATTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CCGTGGCTGTGTCAGTGTCTCCGACAGCAGACGCCCTGCGAGGTGTACGGCTTGT 1449  
1403 TCGGGCTGTCTGCAAGGTGTGTCTGCTAGTCTCTGGGTACAGGCACATGAGCG 1462  
1450 GCGGGCTGTCTGCGCGGTGTGTGCCCCAGGCTCTGGGGCTCCAGCACAAGACG 1509  
1463 CCGCTTTTAAAGACTTAAAGAGTTCTATCTCTGTGGGAAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAAGTTCTATCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1569  
1523 GCAGGAAGTGTGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGAAGATGAGCTGCGGACTGCGCTTGGCTGCGAGAGCCGAG 1629  
1583 GAAGGACCGGTGCCCGCTGCAGACACCGTCTGAGGAGAGGATCTCTGGGTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGCGCGCAGACACCGTCTGCGTAGGAGATCTCTGGCCAGTTCT 1689  
1643 GTTCTGGCTGATGGACACATAGCTGTGACAGTGTAGTGTAGTCAATCTTTTACACAGA 1702  
1690 GCATGGCTGATGAGTGTGTACGTCTGAGTGTGTCAAGTCTTTCTTTTATGTACCGGA 1749  
1703 GAGCACATTCAGAGAACAGGCTCTTCTTACCGTAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACACGTTTCAAGAGAACAGGCTCTTTTCTACCGAAGAGTGTGTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACACTTCAGAGAGTGGGCTACCGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGACGTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTGAGGCTACCCAGGACACTGGCTAGCCATGCCATCTCCAGACTGCGGTTCATCCC 1882  
1870 GGTGAGGAGATCGGGAAGCCAGGCGCGCTGTGAGCTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTCGGGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTGACGGGCTCGGCGGATTTGAACATGAGTACGTGCGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGAGGAGCCAGCATTTACCCAGCGTCTCAAGACTCTCTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGAGCCAGCGTCTCACTCGAGGTTGAAGGCACTGTTCAGCGTCT 2049  
2003 CAATATGAGCGGACAAACATCTCACTTATGGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGGGGGCGGGCGCCCGCGCTCTGGGCTCTGTCTGGGCTGGACGA 2109  
2063 CATCTACAGAGCTGGGGCGCTTTGTGCTGCTGTGCTGTGCTGTGAGACAGACCCAG 2122  
2110 TATCCACAGGGCTGGGCGACCTTCGCTGTGCTGTGCGGGCCAGGACCCCGCTGA 2169  
2123 GATGTACTTTTAAAGCAGATGTGACCGGGGCTATGATGCCATCCCGAGGGTAAGCT 2182  
2170 GCTGTATTGTCAAGTGTGATGTGACGGGCGGTACGACACCTTCCCGCAGGACGCT 2229  
2183 GGTGAGGTTGTGCCAATATGATCAGGCACTCGAGAGACGATCTGTATCCGCAAGTA 2242  
2230 CACGGAGGTATCGCCAGCATCATCA---AACCCAGAAACAGTACTGCGTGGCTCGGTA 2286  
2243 TGCAGTGTCCGGAGATAGCAAGSCCAAGTCCCAAGTCCCTTAGAGACAGGTAC 2302  
2287 TGGGTGTCCAGAGGCGGCCCATGAGGACGTCGCAAGGCTTCAAGAGCCACGTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCTTAAAGCATCTCAGGATTCAGA 2362

2347 TACCTTGCAGACCTCCAGCCGTACATGCGAGAGTTCTGTGGCTCACTGCAG-----GA 2400  
2363 TCCAGTGCATCGAGGAACCTCCGTGTGTCATCGACAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGATGCGTCTCATCGACAGAGTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTGTTGATCTTCTTCCTGCACTCTCTCGCTCACAGTGTCTAAAGATTGGTGA 2482  
2461 CAGTGGCTCTTCGACGCTCTTCTACGCTTCTTCTCATGTGCCACCAAGCCGTGCGCATCAAGGG 2520  
2483 CAGGTGCTATACGACGTGCCAGGCACTCCCGAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAGTCTTACGTCAGTGCAGGGGATCCCGAGGGTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAGAACAGCTGTTTGTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGCAATGAGAGAACAGCTGTTTTCGCGGATTCGCGGGACCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGGTGAAGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCTTGTGTTGAGTATTTCTTGTGTTGACACCTCACTCACCCACGGAAC 2700  
2663 CTTCTCAGCACCTTGTGCTCATGGCTTCTGAGTATGGGTGATGAATACTTGCAGAA 2722  
2701 CTTCTCAGGACCTTGTGCTCGAGGTGCTCTGAGTATGGTGTGCTGAACTTGCAGAA 2760  
2723 GACGTGTGAACTTCCCTGTGAGCTGTGTAACCTGGGTGGTGCAGCTCCATACAGCT 2782  
2761 GACGTGTGAACTTCCCTGTGAGAGACAGGCGCTGGGTGGACGCGCTTTGTTGAGT 2820  
2783 GCTGCTCACTGCTTGTTCCTGTGTGCTTCTGTGACACTCAGACTTTTGGAGT 2842  
2821 GCGGCGCCAGCGCTTATTCCTTGTGCGGCTCTCTGATACCCGACCCCTGGAGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACTTCCAGAG 2902  
2881 GCAGAGCGACTACTCCACTATGCCGAGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTAAAGCTGGGAAGACCATGCGGAACAAGCTCTGTGCGTCTTGGGTGAAAGTG 2962  
2941 CCGCTTCAAGCTGGAGGAACATGCGTGCAGAACTCTTTGGGCTCTTGGGTGAAAGTG 3000  
2963 TACCGTCTATTCTTAGACTTGCAGGTGAAACAGCTCCAGACAGTGTGCATCAATATATA 3022  
3001 TCACAGCTGTGTTCTGAGTTTGCAGGTGAAACAGCTCCAGACGCTGSCACCAACATCTA 3060  
3023 CAGATCTTCTGCTTTCAGGCTACAGGTTCCATGCTGTGATTCAGGTTTCCCTTTGA 3082  
3061 CAGATCTCTGCTGAGGCTACAGGTTTCAAGCATGTGTGCTGAGCTCCCATTTCA 3120  
3083 CCAGCGTGTAGGAAGAACCTCAATCTTTCTGGGATCATCTCCAGGCCAAGCATCTG 3142  
3121 TCAGCAAGTTTGAAGAAACCCACATTTTCTGCGGTCTATCTGTACACGCGCTCCCT 3180  
3143 CTGCTATGCTATCTTGAAGGTCAAGATCCAGGAATGACACTAAG-----GCTC 3193  
3181 CTGCTACTCATCTTGAAGCCAGAACGAGGATGTCGCTGGGGCCAGGCGCGC 3240  
3194 TGGCTCTTCTTCTTGAAGCGGCAATTGGCTCTGTGTCACAGGCTTCTGTCTCAAGCT 3253  
3241 CCGGCTCTGCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATCTCTGTCTCAAGCT 3300  
3254 GGTGCTCATCTGTGCTATCAAAATGTCTCTGGGACTCTGAGGACAGCCCAAACT 3313  
3301 GACTGACACCGGTGTCACTAGTGCCTCTCTGGGTCTACTCAGGACAGCCCAAGCGCA 3360  
3314 GCTGTGCCGGAAGCTCCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGAGCCAGC 3373  
3361 GCTGAGTCGAGAGCTCCCGGAGCGACGCTGACTGCTGCTGGGCGCGCAGCAACCCGCG 3420  
3374 CTTAAGCAGACTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTGTGACTGA 3454

:ch completed: February 2, 2004, 01:50:58  
time : 12479 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: February 1, 2004, 19:29:41 ; Search time 863 Seconds  
(without alignments)  
10935.383 Million cell updates/sec

le: US-09-042-460-1

fect score: 3496

ence: 1 GAATTCGGGTCGGAGGCC.....CGAGCTCGGTACCAAGCTT 3496

ring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

ched: 2552756 seqs, 1349719017 residues

al number of hits satisfying chosen parameters: 5105512

imum DB seq length: 0

imum DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : N Geneseq\_19Jun03.\*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	ID	Description
1	1600.8	45.8	3955	19 AAV22379	Human telomerase r
2	1600.8	45.8	4015	20 AAV230154	cDNA encoding a hu
3	1600.8	45.8	4015	20 AA208150	Human telomerase r
4	1600.8	45.8	4015	20 AA220279	Human telomerase r
5	1600.8	45.8	4015	20 AA200724	Human telomerase c
6	1600.8	45.8	4015	22 AA45901	Human hTERT gene.
7	1600.8	45.8	4015	24 AA46821	Human telomerase r
8	1600.8	45.8	4015	24 ABA97534	Cancer cell discr

9	1600.8	45.8	4015	25 AB222474	Human telomerase r
10	1600.8	45.8	4015	25 AB218391	Group III cDNA can
11	1600.8	45.8	4023	19 AAV60320	Human telomerase g
12	1600.8	45.8	4027	20 AAX89424	Human EST coding
13	1600.8	45.8	4027	21 AAX29388	hEST2, a human tel
14	1600.8	45.8	4042	20 AAV72117	Human catalytic te
15	1600.8	45.8	4070	24 ABL53711	Human telomerase c
16	1599.2	45.7	3798	19 AAV27876	Human telomerase p
17	1596	45.7	3399	24 AB235720	Human telomerase r
18	1596	45.7	3399	24 ABX09963	Human telomerase r
19	1596	45.7	3399	24 ABV78144	Human polynucleoti
20	1596	45.7	3399	24 ABL91685	Human telomerase c
21	1596	45.7	3964	20 AAX18254	Human telomerase c
22	1594.6	45.6	3396	20 AAX18266	Telomerase coding
23	1594.6	45.6	3396	22 AAX44366	Human telomerase n
24	1594.6	45.6	3396	22 AAH48235	Heart muscle cell
25	1594.6	45.6	3396	22 AAH49601	Human coding seque
26	1594.4	45.6	4037	19 AAV22428	Human telomerase r
27	1593.6	45.6	3453	25 AB276217	Human TERT coding
28	1593.6	45.6	13766	24 AAD46790	pgRN145 plasmid DN
29	1584.8	45.3	8742	24 AAD46793	PGRB5A plasmid DNA
30	1526.8	43.7	3918	20 AAX18269	Telomerase coding
31	1526.8	43.7	3918	20 AAX18278	Telomerase coding
32	1508.8	42.4	3203	20 AAX18268	Altered C-terminus
33	1480.6	41.8	7498	21 AAB63786	Telomerase (ver. 2
34	1462.6	41.8	7498	24 ABL50133	Nucleotide sequenc
35	1462.6	41.8	7498	24 ABL50133	Mouse telomerase g
36	1438	41.1	3167	20 AAX18271	Altered C-terminus
37	1438	41.1	3167	20 AAX18280	Altered C-terminus
38	1377.2	39.4	3223	20 AAX18277	Telomerase coding
39	1374.4	39.3	7688	20 AAX18351	Human telomerase r
40	1323.2	37.8	3855	19 AAV22382	Telomerase coding
41	1312.4	37.5	7797	20 AAX18350	Truncated telomera
42	1297.8	37.1	3069	20 AAX18276	Truncated telomera
43	1297.8	37.1	3069	20 AAX18267	Human telomerase p
44	1264.6	36.2	2848	19 AAV27872	Truncated telomera
45	1230.2	35.2	3033	20 AAX18270	Truncated telomera

#### ALIGNMENTS

##### RESULT 1

AAV22379  
ID AAV22379 standard; cDNA; 3955 BP.

XX AAV22379;

XX AC  
XX 13-AUG-1998 (first entry)

XX DT Human telomerase reverse transcriptase encoding cDNA.

XX DE Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis; prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.

XX KW Homo sapiens.

XX OS

XX Key

XX Location/Qualifiers

XX CDS

XX 56..3454

XX /\*tag= a

XX /product= "telomerase reverse transcriptase"

XX GB2317891-A.

XX PD 08-APR-1998.

XX PF 01-OCT-1997; 97GB-0020890.

XX PR 14-AUG-1997; 97US-0915503.

XX PR 01-OCT-1996; 96US-0724643.

XX PR 18-APR-1997; 97US-0844419.

XX PR 25-APR-1997; 97US-0846017.

XX PR 06-MAY-1997; 97US-0851843.







55.

Homo sapiens.

Key Location/Qualifiers

CDS 56..3454

/\*tag= a

/product= "telomerase reverse transcriptase"

W09950392-A1.

07-OCT-1999.

30-MAR-1999; 99WO-US06898.

31-MAR-1998; 98US-0112006.

(GERO-) GERON CORP.

Gaeta FCA;

WPI; 1999-610845/52.

P-PSDB; AA43621.

Eliciting an in vivo immune response for prevention and treatment of cancers

Disclosure; Fig 2; 26pp; English.

The present sequence encodes a human telomerase reverse transcriptase (TRT) polypeptide. The protein is used in the method of the invention. The specification describes a method for activating a T lymphocyte, comprising contacting the T lymphocyte with a dendritic cell that expresses a TRT peptide in the context of a MHC class I or MHC class II molecule. The protein causes induction of an in vivo immunological response to telomerase activity. Cancer cells are characterized by expression of endogenous TRT gene and the presence of detectable telomerase activity. Therefore, by eliciting a specific immune response to TRT or to TRT-expressing cells, it is possible to selectively target proliferating cells for immunological destruction. The method is used for eliciting an in vivo immune response to telomerase by activating a T lymphocyte, and is useful for prevention and treatment of cancers and other proliferation diseases/conditions.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGCGCTTGTAGCAATGACCCGCGCTCTCGTTGCGCCCGCGTGCCTCTCTGTGCG 82

|||||

40 CCGCGCACCCCGCGATGCGCGGCTCCCGCTGCGAGCCGCTCCCTGTGCG 99

|||||

83 CAGCGATACCGGAGGTGTGCGCGCTGCAACCTTTGTGCGCGCTGCGGCCCGAGGG 142

|||||

100 CAGCCACTACCGAGGTGCTCCCGCTGSCACAGTTCTGTGCGCGCTTGGGCCCCAGGG 159

|||||

143 CAGCGCGCTTGTGCAACCCGCGGACCCGAAGATCTACCGCACTTTGTGTGCCAATGCT 202

|||||

160 CTGCGCGCTGTGTGACGCGCGGGAACCGCGCGCTTTCGCGCGCTGTGTGCGCCAGTGCCT 219

|||||

203 AGTGTGCATGCACTGGGCTCACGCTCCACCTGCGACCTTTCTTCCACAGGTGTC 262

|||||

220 GGTGTGCTGCTGGGACGACCGCGCGCCCGCCCGCCCTCTCTTCCGCAAGGTGC 279

|||||

263 ATCCCTGAAGAGCTGGTGGCCAGAGGTTGTGCAGAGACTCTGCGAGCGCAACGAGAGAA 322

|||||

280 CTGCTTGAAGAGCTGGTGGCCCGAGTGTCTGCAGAGGCTGTGCGAGCGCGCGGAAGAA 339

|||||

323 CGTGTGCGCTTTTGGCTTGTAGCTCTTAACGAGCCAGAGCGCGGCTCTCCCATGGCTTT 382

|||||

340 CGTGTGCGCTTTGCGCTCTGCGCTGCGAGCGGCGCCCGCGGGGGCCCCCGAGGCTTT 399

QY 383 CACTAGTAGCGTGTAGCTACTTGCCCAACTGTATTGAGACCTCTGCGTGTGAGTGG 442

Db 400 CACCAACCGGTGCGAGCTACTCTGCCCAACACGCTGACCGACCTGCGGGGAGCGG 459

QY 443 TGCAATGATCTACTGTGAGCGAGTGGCGGACACCTGTGCTGTACCTGCTGTGCA 502

Db 460 GCGTGGGGCTCTGCTGCGCGCGTGGCGGACGAGCTGTGCTTCACTGCTGTGCA 519

QY 503 CTGTGCTCTTTATCTTCTGTGTGCCCCCAGCTGTGCTTACAGAGTGTGTGGTCTCCCT 562

Db 520 CTGCGCGCTCTTTTGTGTGTGCTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 579

QY 563 GTACCAAAATTTGTGCCACCAAGATATCTGCGCTCTGCTGTGCTGTGCTGTGCTGT 622

Db 580 GTACCACTGCGGCTGCTCACTCAGSCCGCGCCCGCCACACGCTGTGGACCCCGA 639

QY 623 CCGACCGTGGGAGGAATTTCACTAACCTTAGTCTTACAAACAGATCAAGAGCAGTAG 682

Db 640 GCGTCTGGGATGC-----GAAACGGCGCTGGAACCATAG 672

QY 683 TCGCAGGAGACCCGAAACCCCTTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742

Db 673 CGTCAGGAGCGCGGTCCCTTGGCTTGCAGCCCGCGGTGCGAGAGCGCGGGG 732

QY 743 TCTCACAGTACAGTGTGCTTCACTAAGAGGCGCAGATGCTATCTCTGTGCTGCGAG 802

Db 733 CAGTCCAGCGAAGTCTGCGCTTGCAGAGGCGCGCTGCGCTGCGCTGCGCTGAGCC 792

QY 803 GGAGGAGGAGCCACAGGCGAGTGTACTACCAACCCCATCAGCAATCATGCTGCTGCA 862

Db 793 GGAGGAGGAGCCCGGTGCGGAGGCTTCTGCGGCGCCACCGCGGAGAGCGCTGAGCG 852

QY 863 TCTGTGCTGCTGCGGAGTGTGCTACT-----GCAGAGAAAGATTTCTTCTTAA 913

Db 853 TGACCGTGTGTTCTGTGTGTGCTCACTGCGAGACCGCGCGAAGACCATCTTTGA 912

QY 914 AGAAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970

Db 913 GGGTGTGCTCTCTGTGCGAGCGGCGCTTCCACCCCATCGTGGGCGCGCGAGCACCG 972

QY 971 CTCACATCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027

Db 973 CCCCCATCCACATCGCGGCGCACCGTCCCTGCGGACACGCTTGTCTCCCGCTGTAC 1032

QY 1028 TGAGACAGACATTTCTTTTCTTCTCAGGAGAGTGTGCGAGAGCGTCTTAAACCTCAT 1087

Db 1033 CGAGACCAAGCACTTCT 1089

QY 1088 CCTACTCAGCAACCTCCAGCTTAACTTGAATGCGGCGAGAGCTGTGCTGCTGCTGCT 1147

Db 1090 CTTACTCAGCTCTCTGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149

QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGACCATCTGCGAGGACACCGCTTATCGGCTG 1207

Db 1150 TCTGGTTCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1209

QY 1208 ATACTGAGAGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267

Db 1210 CTACTGCAAAATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269

QY 1268 TGTGAGACTCTCTCAGGTTCACATTTGCAAGTTCGAGTTTCCAAACAGCAAAACAG 1327

Db 1270 CGGGGTGCTCTCTCAAGAGCGACTGCGCGCTGCGAGTGTGCGTCAACCCAGAGCG 1329

QY 1328 CT-----TGAACACGAGCC 1342

Db 1330 CTGTGCGCGGAGAGCCCGAGGCTCTGTGTGCGCGCGCGCGAGGAGGAGGACAGACCC 1389

QY 1343 ACCGACCTCATGATTTGCTCCGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402

Db 1390 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449

```
1403 TCGGCGCTCTCTCTCAAGGTGGTCTCTGTAGTCTCTGTGGGTACCGAGCACATGAGCG 1462
1450 GCGGCGCTGCTGCGCGCGGTGGTGGCCGCCAGGCTCTGGGGCTTCAGGACACACGACG 1509
1463 CCGCTCTTTTAAAGACTTAAGAAAGTTTCATCTCGTTGGGAAATACGGCAAGCTATCACT 1522
1510 CCGCTCTCTCAGGAACACCAAGAAAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569
1523 CGAAGAACGTGATGTGGAAGATGAAGAGTAGAGATTTGCCACTGGCTCCGACGACGCCGGG 1582
1570 GAGAGAGCTGACGTGGAAGATGAGCGTCCGGGACTGCGCTTGGCTGCGAGGAGCCGAG 1629
1583 GAAGGACCGTGTCCCGCTGTCAGAGACACCGTCTGAGGAGAGGATCTCTGGCTACGTTCCCT 1642
1630 GGTGGCTGTGTCCGCGCGCAGACACCGTCTGCGTGAGGAGATCTTGCCCAAGTTCCCT 1689
1643 GTTCTGGCTGATGGACACATAGTGGTACAGCTGCTTAGGTTCATCTTTTACATCAGAA 1702
1690 GCACTGGCTGATGAGTGTGATCGTCTCGAGCTGCTCAGGTCTTTTCTTTTATGTACGGA 1749
1703 GAGCACATTTCCAGAAAGACAGGCTCTTCTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762
1750 GACCACTTTCAAGAGACAGGCTCTTTTCTA CCGNAGAGTGTCTGTGAGCAAGTTGCA 1809
1763 GAGCAATGGAGTACGCAACACCTTGAAGAGTGGGCTACGGAGCTGTCAACAAGGA 1822
1810 AAGCAATGGAAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTGGAAGCAGA 1869
1823 GGTCAAGCAATCACCAGGACACCTGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCC 1882
1870 GGTCAAGCAATCAGGAGCAGCAAGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCC 1929
1883 CAAAGCCCAACCGCCCTCGGCGCATTTGTAACATGAGTTATAGCATGGGTACCGAGCTTT 1942
1930 CAAAGCTCAGCGGCTCGCGCGCATTTGTAACATGAGTACTGCTGCGGAGCCAGAACGTT 1989
1943 GGGCAGAGGAGCAGGCGCCAGCTTCAACCGAGCTCAAGACTCTCTTCAAGCTGCT 2002
1990 CCGCAGAGAAAGGGCGGAGCTCTCACCCTCAGGGTGAAGCACTGTTCAGCGGTCT 2049
2003 CAACTATGAGCGGACAAACATCTCTCACTTATGGGTCTTCTGTA TGGGTATGAATGA 2062
2050 CAACTACGAGCGGCGCGCGCCCGGCTCTCTGGGCGCTCTGTGTGGCTTGACGA 2109
2063 CATCTACAGGACTCGGCGGCTTTGTGCTCGGTGGGTGCTCTGACACAGACACCCAG 2122
2110 TATCCACAGGCGCTCGCGCACTTCTGCTGCTGCTGTGCGGGCCAGGACCGCGCGCTGA 2169
2123 GATGTACTTTTGTAAAGGAGATGTGACCGGGCTTATGATGCCATCCCCAGGCTAAGCT 2182
2170 GCTGTACTTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCAGGACAGGCT 2229
2183 GGTGAGGTTTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242
2230 CACGGAGGTCTATCGGCAGCATCATCA---AACCACAGAAACAGTACTCGTGGTGGTA 2286
2243 TGCAGTGTGCGGAGATAGCCAGGCGCAAGTCCAGTCTTATAGGAGACAGGTAC 2302
2287 TGCCGTGTGCAAGAGCGGCCATGGGCACTCCGCAAGGCTTCAAGAGCCACGCTCTC 2346
2303 CACCTCTCTGACCTCCAGCCATATACGGGCAAGTCTTAAAGCATCTGAGGATTCAGA 2362
2347 TACCTTGACAGACCTCCAGCGGTACATCGACAGTTCGTGGCTCACCTGCAG-----GA 2400
2363 TGCCAGTGCATGAGAACTCGGTGTGATCGAGCAGAGCATCTCTATGATGAGAGCAG 2422
2401 GACCAAGCCGCTGAGGGATGCGGTGCTCATCGAGCAGAGCTCTCCCTGATGAGCCAG 2460
2423 CAGCAGCTTTTGAATCTTCTTCTGCACTTCTCGTCAAGTGTCTGTAAGATTTGGTGA 2482
2461 CAGTGGCTCTTCGAGGCTTCTTCTTACGCTTTCATGTGCCACCAAGCGGTGGCATCAGGG 2520
2483 CAGGTGCTATACCGAGTGCAGGGCATCCCCCAGGGCTCCAGCCATTCACACCTGTCTGTG 2542
2521 CAAGTCTTACGTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCAGCGTCTCTG 2580
2543 CAGTCTGTGTTCCGAGACATGAGAACAAAGCTGTTTGTGTAGGTGCAGCGGATGGTT 2602
2581 CAGCTGTGCTACCGGACATGAGAACAAAGCTGTTTCCGGGATTTCCGGCGACCGGCT 2640
2603 GCTTTTACGTTTGTGATGACTTCTGTTGGTGAAGCTCACTTGGACCAAGCAAAAC 2662
2641 GCTCTGTGTTTGTGTGATGATTTCTTTTGGTGAACACCTCACCTCACCCACGCAAAAC 2700
2663 CTTCTCAGCACCCCTGTCTCATGGCTTCTTCTGATGATGGGTGATGATATAAATTGAGAA 2722
2701 CTTCTCAGGACCCCTGTCTCGAGGTGCTCCCTGATGATGGTGGTGAATCTTGGAA 2760
2723 GACAGTGGTGAACCTTCCCTGTGAGGCTGTATCCCTGGGTGGTGGAGCTCCATACAGCT 2782
2761 GACAGTGGTGAACCTTCCCTGTGAGAACGAGGGCTTGGGTGGCAGCGCTTTTGTTCAGAT 2820
2783 GCCTGCTCAGCTGCTGTTTCCCTGGTGGCTTGTCTGTGGACACTCAGACTTTGGAGCT 2842
2821 GCCGCGCCACGGGCTATTCCCTGGTGGGCTGTCTGTGGATACCCGGAACCTGGAGGT 2880
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACCTTCAGAG 2902
2881 GCAGAGGAGTACTACTCCAGTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940
2903 TGTCTTCAAGCTGGGAGACCATGCGGAACAGCTCCTGCTGGTCTTGGGTGGAAGTG 2962
2941 CGGCTTCAAGCTGGGAGAACATGCTGCCAACTCTTTGGGTCTTGGGCTGAAGTG 3000
2963 TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCATCAATATATA 3022
3001 TCACAGCTGTTTCTGATTTGCAAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA 3060
3023 CAAGATCTTCTGTCTCAGGCTACAGGTTCATGATGATGATTCAGCTTCCCTTTGA 3082
3061 CAAGATCTTCTGTCTCAGGCTACAGGTTTACGATGTGTGCTGCTGCTCCCATTTCA 3120
3083 CCAGCTGTTTAGGAAGAACCTCACATTTCTTCTGGGCATCATCTCCAGCCAAAGCATCTG 3142
3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTTCTGGCGTCACTCTGACACGCGCTCCCT 3180
3143 CTGTATGCTATCTCTGAAGTCAAGATCCAGAAATGACATAAG-----GCCTC 3193
3181 CTGTACTCTCATCTCTGAAGCCAAAGAACGAGGGATGTCTGGGGGCCAAGGGCGCCGC 3240
3194 TGGCTCTTCTCTCTCAAGCCGACATTTGGCTCTCTACAGGCTTCTCTGCTCAAGCT 3253
3241 CGGCGCTCTGCGCTCCAGGCGCTGAGTGGCTGTGCTGCCACCAAGCATTCCTGCTCAGCT 3300
3254 GGCTGCTCATTTCTGTCTATCTAATAATGTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313
3301 GACTCGACACCGTGTCACTACCTGCGACCTCTCTGGGCTCACTCAGGACAGCCAGACGCA 3360
3314 GCTGTGCGGAGAGCTCCAGAGGGGCAATGACCATCTTAAAGCTGAGCTGAGCCAG 3373
3361 GCTGAGTGGAGAGTCCCGGGAGCAGCTGATGCTGCCCTGGAGCGCGCAGCCACCCGCG 3420
3374 CCTAAGCACAGACTTTTCAGACCATTTTGGACTAA 3407
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

RESULT 3
AAZ08150
ID AAZ08150 standard; cDNA; 4015 BP.
XX
AC AAZ08150;
XX
DT 17-JAN-2000. (first entry)
XX
DE Human telomerase reverse transcriptase cDNA.
```

ue Feb 3 10:49:29 2004

Human telomerase reverse transcriptase; hTERT; telomerase; hEST2; catalytic protein component; cell proliferative capacity; cell immortality; neoplastic phenotype; diagnostic application; prognostic application; telomerase related condition; cancer; therapeutic agent; telomerase expression; telomerase activity; ds.

Key	Location/Qualifiers
CDS	56..3454 /*tag= a /product= "Human telomerase reverse transcriptase" /transl_except= (pos:1877..1879, aa:Gln)

· WO9950279-A1.

07-OCT-1999.

31-MAR-1999: 99WO-US07160.

31-MAR-1998: 98US-0052919.

(GERO-) GERON CORP.  
(UYTE-) UNIV TECHNOLOGY CORP.

Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB, Andrews WH;

WPI; 1999-610834/52.  
P-PSDB: AAY28881.

Antisense polynucleotides for human telomerase reverse transcriptase used for diagnosing or treating cancer -

Claim 1: Fig 1: 31pp: English.

The present sequence encodes for human telomerase reverse transcriptase (hTERT). This is the catalytic protein component of telomerase and is also referred to as hstr2. This correlates with cell proliferative capacity, cell immortality, and the development of a neoplastic phenotype. Human TTR antisense oligonucleotides are useful for diagnostic or prognostic applications to telomerase related conditions, including cancer. They are also useful as therapeutic agents, for inhibition of telomerase expression and activity.

Sequence 4015 BP: 663 A; 1364 C; 1274 G; 714 T; 0 other;

Query Match	45.8%	Score 1600.8	DB 20	Length 4015
Best Local Similarity	68.9%	Pred. No. 0		
Matches 2379	Conservative	0	Mismatches 967	Indels 108
				Gaps 9
23	CCCGGCGCTTGAGACAAATGACCCGCGCTCCCTGCTTGCCCGCGGTGCGCTCTCTGCTGCG	82		
40	CCGGCCACCCCGCGATGCGCGGCTCCCGCTGCCGAGCGTGGCTGCTGCTGCG	99		
83	CAGCCGATACCGGAGCTGTGCCCTGGCAACTTTGTGGCGGCGCTGGGCGCCGAGG	142		
100	CAGCCACTACCGGAGGTCTCCGCTGCGCACGTTCTGTGGGCGCTGTGGGCCCCAGG	159		
143	CAGCGGCTTGTGCACACCGGGGACCCGAGATCTACCGCACTTTGGTTGCCCAATGCCT	202		
160	CTGGCGGCTGTTGACGCGCGGGACCCGCGGCGCTTTCGCGCGCTGTGTGGCCAGTGCCT	219		
203	AGTGTGCATCACTGGGGCTCACAGCTCCACTTCCGCGACCTTTTCCTTCCACCAAGTGTTC	262		
220	GGTGTGCTGCCCTTGGGACGACGCGCGCCCCCGCGCCCTCTTCTTCGCGCAGTGTTC	279		
263	ATCCCTGAAGAGCTGGTGGCCAGGTTGTGCAGAGACTCTTGCAGGCGCACGAGAGAA	322		
280	CTGCCTGAAGAGCTGTGTGGCCCGAGTGTGTGCAGAGGCTGTGCGAGGCGCGCGCGAAGA	339		
323	CGTGTGCTTTTGGCTTTTGAGCTGTGCTTAAACGAGGCCAGAGCGGGGCTCTCCCATGGCCCTT	382		





23	CCGGCGCTTGAGACAAATGACCCGGCGCTCCTGTTGCCCGCGGTGCGCTCTCTGCTGGC	82
40	CCGGGCACCCCGCGATGCGCGCGCTCCCGCTCCGAGCGTGTGCTCCTCTGCTGGC	99
83	CAGCCGATACCGGAGGTGTGGCGCTGSCAACCTTTGTGCGCGCGCTCGGCGCCCGAGGG	142
100	CAGCCACTACCGGAGGTGTGCCGTGGCCACGTTCTGTGCGGCGCTTGGGCGCCAGGG	159
143	CAGCGCGCTTGTGCACCCGGGAGCCGAGATCTACCGACACTTTGGTTGCCCAATGCT	202
160	CTGGCGCTGGTGCAGCGGGGACCCGCGGCTTTCGCGGCTGTGTGGCCAGTGCT	219
203	AGTGTGATGCACTGGGGCTCAGAGCTCCACTGCGGACCTTTCCTTCCACCGAGTGTC	262
220	GGTGTGCTGTCCTTGGGACGACGGGCGCCCCCGCGCCCCCTCTCTCGCGAGGTGTC	279
263	ATCCCTGAAGAGCTGGTGGGCCAGGGTTGTGACAGACTCTGGGCGGCGCAACGAGAGAA	322

1343 ACCGACCTCATGATTTGCTCCGCTGCAACAGCATCCTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CCGTCGCTTGTGAGCTGTCTCGCCAGACACAGCAGCCCTCGCAGGTGTACGGTTCGT 1449  
1403 TCGGGCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACAGGCAAAATGAGCG 1462  
1450 GCGGCTGTCTCGCCGCTGTGTGCCCCAGGCTCTTGGGCTCCAGGCACAGCAAG 1509  
1463 CCGCTTCTTAAGAACTTAAGAACTTATCTGTTGGGAAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAACTTATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGAACTGATGTGGAAGATGAAGATGAGAGATGCCCATGCTCCGAGCAGCCGCG 1582  
1570 GCAGAGCTGACGTGGAAGATGAGCTGCGGACTGCGCTTGGCTGCGAGAGCCCAAG 1629  
1583 GAAGGACCGTGTCCCGCTGCAAGACACCGTGTGAGGAGAGGATCTGGTATCTCT 1642  
1630 GGTGGCTGTGTCCGCGCGCAGACACCGTCTGCGTGAGGAGATCTTGGCCAAAGTTCCT 1689  
1643 GTTCTGCTGATGGACACATACCTGCTGACAGTGTAGTCAATCTTTTATACACAGA 1702  
1690 GCATGCTGATGAGTGTGACGTGCTGAGCTGCTCAGGTCTTCTTTTATGTCAACGA 1749  
1703 GAGCACATTCAGAAAGAAAGGCTCTTTCTTACCCTAAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAAGAAAGGCTCTTTTCTACCGGAAGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACACCTTGAGAGTGTGCGCTACCGGAGCTGTACAGAGGA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGTGACGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTGAGGATCAACAGACACCTTGGGTAGCCATGCCCCTCTGACAGCTCGCTTCAATCCC 1882  
1870 GGTGAGGACATCGGAAGCAGGCGCGCCCTGTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGCGCTGCGGCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTTGAAGGCTGCGGCCGATTTGTGAACATGAGTACGTGCTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAAAGAGCAGGCCAGCATTTCAACAGGCTCTCAAGACTCTCTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCCGAGGCTCTACCTCGAGGTGAAGGCACTGTCAGGCTGCT 2049  
2003 CAATATGAGCGCAAAACATCCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATAGAGCGGCGCGCGCCCGGCTCTGCGGCTCTGTGCTGGGCTTGGACGA 2109  
2063 CATCTAGGACCTGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCAGAGGCTTGGCGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATGATCTTTTAAGCAGATGTGACCGGGGCTATGATGCCATCCCCAGGGTAAAGCT 2182  
2170 GCTGATCTTTTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTGTGCTCAATATGATCAGGCTCGGAGAGCAGTACTGTATCCGCGAGTA 2242  
2230 CAGGAGGTATGCGCCAGCATATCA---AACCCAGAACACGTAATGCTGCTGCTGCTGCTGCT 2286  
2243 TGCAGTGGTCCGAGAGATAGCAAGGCCAAGTCCCAAGTCTTTTGGAGACAGGTCAAC 2302  
2287 TGCGTGTTCAGAGAGCCCGCATGGGCACTGCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCAGCCATATAGGCGCAATGCTTTAAGCATGTGCAAGATTCAGA 2362  
2347 TACCTTGACACAGCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGTCACTGAGGAACCTCGTGTGTCGAGAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCTCCTGATGAGGCCAG 2460

QY 2423 CAGCAGCCTGTTTGACTTCTCTCTGCACTTCTGCGGTACAGTGTCTGTAAAGATTGGTGA 2482  
Db 2461 CAGTGGCTCTTTCGACGCTTCTCTACGCTTATGTGCGCACACGCGGTGCGCATCAGGG 2520  
QY 2483 CAGGTCTATATACGAGTGCAGGCGATCCCCAGGGTCCAGGCTATCCAGCCTGCTCTG 2542  
Db 2521 CAACTCTACGTCCAGTCCAGGGATCCCGAGGGCTCCATCTCTCCAGCCTGCTCTG 2580  
QY 2543 CAGTCTGTCTTTCGAGAGATGAGAAACAAGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
Db 2581 CAGCTGTGCTACGCGCACATGGAGAAACAAGCTGTTTCGGGGATTCGCGGGGACGGGCT 2640  
QY 2603 GCTTTTACGCTTGTGATGACTTCTGTGTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
Db 2641 GCTCCTGCTTGTGATGATTTCTTGTGTGACACTCACTCACCACCGGAAAC 2700  
QY 2663 CTTCTCAACAACCTTGGTCCATGCGGTTCTGAGTATGGGTGACATGATAAACTTGCAGAA 2722  
Db 2701 CTTCTCAGGACCTTGGTCCGAGGTGCTCCTGAGTATGGCTGCGTGTGAACCTTTCGGAA 2760  
QY 2723 GACAGTGTGAACCTTCCCTGCGAGCCTGATACCTGGTGGTGCAGCTCCATACCACT 2782  
Db 2761 GACAGTGTGAACCTTCCCTGTAGAAAGCAGGCCCTGGGTGGCAGCGCTTTTGTTCAGAT 2820  
QY 2783 GCTGCTCACTGCTGTTTCCCTGCTGCTGGCTTGTGCTGGACACTCAGACTTTTGGAGGT 2842  
Db 2821 GCGGCGCCACGCGCTATTCCCTGCTGGGCTGCTGCTGATACCGGACCCCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTATGCCAGACCTCAATTAAGACGAGCCTCAGCTTCCAGAG 2902  
Db 2881 GCAGCGGACTACTCCAGCTATGCCCGGACCTCCATCAGACCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAGAGTGGGAAGACCATCGGAAACAAGCTCCTGCTGGTCTTTCGGTTTGAAGTG 2962  
Db 2941 CGGCTTCAAGGCTGGAGGAAACATCGTGGCAAACTCTTTCGGGCTCTTTCGGCTGAAGTG 3000  
QY 2963 TACGCTGTATTTCTAGACTTTCAGGTGAAACAGCTCCAGACAGTCTGCAATCAATATA 3022  
Db 3001 TCACAGCTGTTTCTGGATTTTCAGGTGAAACAGCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAAGATCTTCTGCTTCCAGGCTACAGGTTCATGATGTGATTCAGCTTCCAGCTTCCCTTTGA 3082  
Db 3061 CAAGATCTTCTGCTGAGGCTACAGGTTTACGCAATGTGTGCTGAGCTTCCCATTTCA 3120  
QY 3083 CAGCGTGTATGGAAGAACCTCAATTTCTTGTGGCATCATCTCCAGCCAAAGCATCTG 3142  
Db 3121 TCAGCAAGTTTGGAAAGAACCCCACTTTTCTGCGCTCATCTCTGACACGGCTCCCT 3180  
QY 3143 CTGCTATGCTATCTTGAAGTCAAGATCCAGGATGACACTAAG-----GCCTC 3193  
Db 3181 CTGCTACTCCATCTTGAAGCCAGGACGAGGATGCTGCTGGGGCCAAAGGGCGCGC 3240  
QY 3194 TGGCTCTTTTCTTCCAGGCGCACATTTGGCTCTGTCTACAGGCTTTCCTGCTCAAGCT 3253  
Db 3241 CCGGCTCTGCTCCGAGGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
QY 3254 GGTGCTCATTTCTGCTCATCAAAATGCTCTCTGGACCTCTGAGGACAGCCGAAACT 3313  
Db 3301 GACTGACACCGGTGTCACTAGTCCCACTCTGGGTCACTCAGGACAGCCGACGCA 3360  
QY 3314 GCTGTGCCGGAAGCTCCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTCCGAGAGCTCCCGGGACGACGCTGACTGCCCTGAGGCGCGCAGCAACCCGCG 3420  
QY 3374 CTTAAGCAGACTTTTCAGACCATTTTTCAGCTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCGGACTGA 3454

RESULT 5  
AAZ00724  
ID - AAZ00724 standard; DNA; 4015 BP.  
XX

AAZ00724;

06-OCT-1999 (first entry)

Human telomerase catalytic domain DNA.

Telomerase; catalytic domain; human; quantitation; tumour cell; melanoma; body fluid; metastases; T-cell lymphoblastoma; chronic myeloid leukemia; acute lymphatic leukemia; melanoma; pulmonary carcinoma; colon cancer; breast cancer; ss.

Homo sapiens.

DE19804372-A1.

05-AUG-1999.

04-FEB-1998; 98DE-1004372.

04-FEB-1998; 98DE-1004372.

(DAHM/) DAHM M W.

Dahm MW;

WPI; 1999-431408/37.

Quantifying tumor cells by amplifying mRNA encoding the catalytic subunit of telomerase

Example; Fig 1A-B; 26pp; German.

This invention describes a novel method for the quantitation of tumour cells in a body fluid which comprises (1) enrichment or isolation of tumour cells in the sample, (2) amplification of mRNA from these cells that encodes the catalytic subunit of telomerase and (3) quantifying the amount of amplified mRNA. The method is applied to tumour cells derived from (micro)metastases, e.g. associated with a wide range of tumours such as T-cell lymphoblastoma, chronic myeloid or acute lymphatic leukemia, melanoma, pulmonary carcinoma, cancer of colon or breast etc. This sequence encodes a human telomerase protein catalytic domain.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGCGCTTGACACATGACCCCGCTCTCGTTGCCCGCGGTGCGTCTCTGTCGG 82

40 CCCGCGCACCCCGCGATGCCGCGCGCTCCCGCGTCCGAGCGCTCGCTCCGTCGG 99

83 CAGCCGATACCGGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGGCCGAGGG 142

100 CAGCCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGGCCGAGGG 159

143 CAGCGCGCTTGTGMAACCCGGGACCCGAGATCTACCGACATTTGGTTGCCCAATGCCCT 202

160 CTGGCGGCTGTGTGACGCGGGGACCCGGCGGCTTTCCGCGCGCTGTGGTCCCGAGTGCCT 219

203 AGTGTGATGACATGGGGCTCAGAGCTTCCAGCTTCCCTTCCACACAGGTGTC 262

220 GGTGTGCTGTGCTTGGGACGACGGCGCGCCCGCGCCCTCTCTTCGCGCAGGTGTC 279

263 ATCCCTGAAGAGCTGTGGCCAGGGTTGTGACAGATCTTGGAGCGCAACGAGAGAA 322

280 CTGGCTGAAGAGCTGTGGCCCGAGGTGCTCAGAGGTGTGCGAGCGCGCGCGAAGAA 339

323 CGTGTGCTTTTGGCTTTTGTAGCTGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTTT 382

340 CGTGTGCGCTTCGCGCTTCGCGCTGCTGAGCGGGGCGCGGGGGCCCCCGAGGCTTT 399

QY 383 CACTAGTAGCGTGTAGTACTTGGCCCAACAATCTGTTATTGAGACCTTGGCTGTCACTGG 442

DB 400 CACCACACGCTGCGCAGCTACTGCTCCCAACAACGCTGACCGACGCACTGCGGGGAGCGG 459

QY 443 TGCATGGATGCTACTGTTGAGCCGAGTGGGCGACACCTGTGTTGTTCTACCTGTCTGGCACA 502

DB 460 GGGTGGGGGCTGCTGCTGCGCGCGTGGCGACACGCTGCTGGTTACCTGTCTGGCAGC 519

QY 503 CTGTGCTCTTTATCTTCTGGTGCCTCCCGCAGCTGTGCTTACAGAGTGTGTGGTCTCCCT 562

DB 520 CTGCGGCTCTTTGTGCTGGTCTCCAGCTGTGCTTACAGAGTGTGCGGGCGCGCT 579

QY 563 GTACCAATTTGTGCCACCAAGATATCTGGCCCTCTGTGCTCGCTAGTTACAGGCCAC 622

DB 580 GTACCAAGCTGCGCTGCGCTGCGCTGCGCCCGCGCCGCGCACGCTAGTGGACCCGAG 639

QY 623 CCGACCCGTGGCAGGAATTTCACTAACTTAGTGTCTTAAACAAGATCAAGAGCAGTAG 682

DB 640 GCGTCTGGGATGC-----GAAACGGGCTTGGAAACCATAG 672

QY 683 TCGCCAGGAGCACCGAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATCTCAG 742

DB 673 CGTCAGGAGGCGCGGGTCCCCCTGGGCTGCGCCCGGGTGGAGGAGCGCGGGG 732

QY 743 TCTCACCAGTACAAGTGTGCTTTCAGCTAAGAGCCAGATGCTATCTCTGCCGAGAGT 802

DB 733 CAGTGCACGCGGAAGTCTGCGCTTGCCCAAGAGGCCAGGCGTGGCGCTGCCCTTGAGCC 792

QY 803 GGAGGAGGACCCACAGGAGGTGTACCAACCCCATCAGGCAATCATGGTGCACAG 862

DB 793 GGAGCGGACGCGCGTGTGGCGGGTCTTGGGCCCAACCCGGGCGAGGACGCTGTGACCGAG 852

QY 863 TCCTGTCTGCTCCCGCGAGTGCCTACT-----GCAGAGAAAGATTTGTCTTTCAA 913

DB 853 TGACCGTGTGTTCTGTGTGTGTCACTTGCAGACCCGCGGAGAGCCACCTCTTTTGA 912

QY 914 AGAAAGGTGTGTGACTGAGTCTCTC---TGCGTGGTGTGCTGTAAACAAGCCAG 970

DB 913 GGGTGGCTCTCTGGCACGCGCACTTCCACCCATCTCGTGGGCGCCAGCACCGCGG 972

QY 971 CTCACAT---CTCTGCTGTACACACCCCGCCAAATGCTTTTCAGCTCAGGCCATTTAT 1027

DB 973 CCCCCCATCCACATCGCGGCCACCGTCTCTGGGACAGCGCTTGTCCCCCGGTGTACGC 1032

QY 1028 TGAGACAGACATTTCTTTTACTCCAGGGGAGATGGCCAAAGAGCTCTTAAACCCCTATT 1087

DB 1033 CGAGACCAAGCACTTCTCTACTCTCTCAAGCGA---CAAGAGGAGAGCTGCGGCCCTCCT 1089

QY 1088 CTTACTCAGCAACCTCCAGCTTAATCTGACTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147

DB 1090 CTTACTCAGCTCTTGAGGCCACAGCTGACTGCGCTCGAGGCTCTGTGGAGACCATCTT 1149

QY 1148 TCTGGGCTCAAGGCTTAGACATCAGGACCACTCTCAGGACACACCTCTATCGCGTCG 1207

DB 1150 TCTGGGTTCCAGGCCCTGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCCAGCG 1209

QY 1208 ATACTGCGAGATGGGCGCTCTTCCACAGCTGTGTGACCATGACAGTGCAGATGCCATA 1267

DB 1210 CTACTGCAAAATCGGCGCTGTTTCTGAGTGTCTGGAAACACGCGCAGTGTCCCCCTA 1269

QY 1268 TGTCAAGATCTCAGGTACATTTGCAAGTTCGAAACAGCAACCAACAGGTGACAGATGC 1327

DB 1270 CGGGGTCTCTCAAGACGACATGCCCGCTGCGAGTGTGCGGTCAACCCAGACGCGGTGT 1329

QY 1328 CT-----TGAAACACAGCCC 1342

DB 1330 CTGTGCCCGGAGAACCCCGAGGCTCTTGTGGCGGCCCGCAGGAGGAGGACACAGACCC 1389

QY 1343 ACCGACCTCATGATTTGCTCGCGCTGCAAGAGTCCCTGCGAGGTATATGTTTCT 1402

DB 1390 CGTGTGCTGTGACGCTGCTCGCGCAGCAGCAGGCGCCCTGCGAGGTGTACGGCTTGT 1449

QY 1403 TCGGCGCTGTCTCTGCAAGGTTGTGTCTAGTCTCTGGGGTACACAGGCAATGAGCG 1462

1450 GCGGGCTGCTGCGCGGTGCTGCGCCAGGCTCTGGGGCTCCAGGCAACG 1509  
1463 CCGCTTCTTTAAGAACTTAAAGAACTTCACTCTCGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAACTTCACTCTCGTTGGGGAATACGGCAAGCTATCACT 1569  
1523 GCAGGAATCATGTGGAAGATGAAATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629  
1583 GAAGGACCGTGTCCCGCTCCAGAGCAACGCTCTGAGGAGAGATCCTGGTCACTGCTCT 1642  
1630 GCTTGGCTGTGTTCCGCGCGAGAGACGCTCTGCTGAGGAGATCCTGGCCAGTCTCT 1689  
1643 GTTCTGGCTGATGAGACACATACGTGGTACAGCTGCTTGGTCACTCTTTTATCATACAGA 1702  
1690 GCATGGCTGATGAGTGTACGCTCGTGGAGCTGCTCAGGCTCTTCTTTATGTCACGGA 1749  
1703 GAGCACATTCAGAGAAACAGGCTCTTCTTACCTGTAAGAGTGTGGAGCAAGTCA 1762  
1750 GACACGTTTCAAGAAACAGGCTCTTCTTACCTGTAAGAGTGTGGAGCAAGTCA 1809  
1763 GAGCATTTGAGTCAAGCAACCTTGGAGAGTGCCTGTAAGAGTGTGGAGCAAGTCA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTCTCGGAAGCAGA 1869  
1823 GGTGAGCATCACAGACACCTGGCTAGCCATCTGCAAGCTGCGCTTATCC 1882  
1870 GGTGAGGAGCATCGGAAGCCAGGCGCGCTGCTGAGCTGCGGAGCTCTCGGAAGCAGA 1929  
1883 CAAGCCCAACCGCTGCGGCGCTTGTGAACATGAGTATAGCATGGTACAGAGCTTT 1942  
1930 CAAGCTTGAAGGCTGCGGCGCTTGTGAACATGAGTATAGCATGGTACAGAGCTTT 1989  
1943 GGGCAGAGGAGCAGGCGCGCTTGTGAACATGAGTATAGCATGGTACAGAGCTTT 2002  
1990 CCGCAGAGAAAGAGGCGCGCTTGTGAACATGAGTATAGCATGGTACAGAGCTTT 2049  
2003 CAATATGAGCGGCAAAACATCTCACTTATGAGGCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGGCGCGGCGCGCTTGTGAACATGAGTATAGCATGGTACAGAGCTTT 2109  
2063 CATCTACAGGAGCTGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCAGAGGCTGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATCTACTTTTGAAGCAGATGTGCGGCGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCT 2182  
2170 GCTGTACTTTTGAAGTGTGATGCGGCGCTTATGATGCTGCTGCTGCTGCTGCTGCTGCT 2229  
2183 GGTGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2242  
2230 CACGAGGCTCATCGCCAGCATCA...AACCCAGAACATGCTGCTGCTGCTGCTGCTGCT 2286  
2243 TGCACTGCTGCGGAGATAGCAAGCCAGTTCACAGTCTCTTTAGCAGACAGTCA 2302  
2287 TGCCGTGCTCAGAGGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346  
2303 CACCTCTCTGACCTCAGGCTATACATGGGCGCTTCTTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTCGAG...GA 2400  
2363 TGCCAGTGCATGAGGAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422  
2401 GACCAGCCCGCTGAGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
2423 CAGCAGGCTGTTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482  
2461 CAGTGGCTCTTTCGACGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGTGTCTATACGAGTGCAGGCGCTTCCCGGAGGCTGAGCTTATCCACCTGCTGCTGCT 2542

2521 CAACTCTACGTCCAGTCCAGGGATCCGAGGGCTCCATCTCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAACAAGCTGTTTCTGAGGTGAGGGGATGSGTT 2602  
2581 CAGCTGTGCTACGGCGACATGAGAAACAAGCTGTTTTCGGGGGATTCGGCGGACGGGCT 2640  
2603 GCTTTTACGTTTGTGATGATTTCTGTTGGTGAAGCTCCTGTTGACCAAGCAAAAC 2662  
2641 GCTCTGCTGCTTGGTGGATGATTTCTTGTGTTGATGATTTCTTGTGTTGATGATTTCT 2700  
2663 CTTCTCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
2701 CTTCTCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2723 GACAGTGTGAATCTTCCCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782  
2761 GACAGTGTGAATCTTCCCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTGCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCACGCGCTTATCCCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGTGATCTCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2902  
2881 GCAGAGGCTGCTTCCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
2903 TGTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
2963 TCAAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3022  
3001 TCACAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060  
3023 CAAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082  
3061 CAAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3083 CCAAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
3121 TCAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
3143 CTGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3193  
3181 CTGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240  
3194 TGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
3241 CGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGTGCT 3313  
3301 GACTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCT 3373  
3361 GCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CTTAAGCAGACCTTTCAGACCAATTTTGGACTAA 3407  
3421 ACTGCTCTCAGACTTCAAGACCATCTCTGACTGA 3454

## RESULT 6

AAH45901 standard; DNA; 4015 BP.

ID AAH45901;

AC AAH45901;

DT 06-SEP-2001 (first entry)

XX Human hTERT gene.

DE

XX

3213..3350

Query Match	45.8%; Score 1600.8; DB 22; Length 4015;
Best Local Similarity	68.9%; Pred. No. 0;
Matches 2379; Conservative	0; Mismatches 967; Indels 108; Gaps 9;
QY	23 CCGCGCCTTGAGCAATGACCGCGCTCTCGTTGCCCGCGGTGCGCTCTCTGCTGCG 82
DB	
QY	40 CCGGCGCACCCCGCGATGCGCGCGTCCCGCTGCCAGCGTGGCTCCCTGCTGCG 99
DB	
QY	83 CAGCCGATACCGGAGAGTGTGGCGCTGGCAACTTTGTGCGCGCGCTGGGGCCCGAGGG 142
DB	
QY	100 CAGCCACTACCGCGAGGTGTGGCGCTGGCACGTTCTGTGCGCGCGCTGGGGCCCGAGGG 159
DB	
QY	143 CAGGCGGCTGTGTGCAACCGCGGAGCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCT 202
DB	
QY	160 CTGGCGGCTGTGTGAGCGCGCGGACCGCGCGGCTTTTCGCGCGCTGGTGGGCCAGTGCT 219
DB	
QY	203 AGTGTGCATGCACCTGGGGCTCACAGCCTCCACCTGCCGACCTTTCTCTCCACCGAGGTGC 262
DB	
QY	220 GGTGTGCGTGCCTTGGAGCGCACGGCGCGCCCGCGCGCCCTCTCTTCGCCCAGGTGTC 279
DB	
QY	263 ATCTCTGAAAGAGCTGTGTGCCAGGFTGTGCAGAGACTGTGCGAGCGGCACGAGAGAAA 322
DB	
QY	280 CTGCTCTGAAGAGAGCTGTGGCCCGAGTGTGTGCAGAGGCTGTGCGAGCGCGCGCGAAGAA 339
DB	
QY	323 CGTGTCTGCTTTTGGCTTTGAGCTGTTTAAAGAGGCCAGAGCGGGCGCTCCCATGCGCTT 382
DB	
QY	340 CGTGTCTGCGCTTCGCGCTTCGTGTGAGCGGGGCCGCGGGGGCCCCCGAGGCGCTT 399
DB	
QY	383 CACTAGTAGCGTGGGTAGCTACTTGGCCCAACACTGTATTGAGACCTGCGTGTGCAGTGG 442
DB	
QY	400 CACACACAGCTGTGCGAGCTACCTGCCCAACACGGTGACCGAGCGCACTGCGGGGAGCGG 459
DB	
QY	443 TGCATGGAATGCTACTGTGTGAGCGAGTGGGCGAGCACTGTCTGTGCTTACCTGCTGGCACA 502
DB	
QY	460 GCGCTGGGGGCTGTGTGTGCGCGCTGTGGCGAGCACTGCTGGTTTCACTCTGCTGGCAG 519

503 CTGTGCTCTTTATCTTCTGCTGCCCGCCAGCTGTGCTACCGAGTGTGTGGCTCTCCCT 562  
520 CTGGCGCTCTTTGTGTGTGTGCTCCAGCTCGCTCTACCGAGTGTGTGGCGCCCGCT 579  
563 GTACCAAAATTGTGCCACCGATATCTGGCCCTCTGTGTCCGCTAGTTTACAGCCCAAC 622  
580 GTACCACTGGCGCTCCACTCAGGCCCGGCCCGCCACACGCTAGTGGACCCCGAAG 639  
623 CCGACCGTGGCGAGGAATTTCACTAACTTAGTTCTTTACAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGGCTGGAACCATAG 672  
683 TCGCCAGGAACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 CGTCAGGAGCGCGGGTCCCTCTGGGCTGCCAGCCCGGGTGGAGGAGCGCGGGG 732  
743 TCTCACAGTACAGTGTGCTTCAGTAAAGAGGCGAGATGCTATCTGTCTCCAGAGT 802  
733 CAGTGCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCCGAGCGTGGCGCTGGCCCTGAGCC 792  
803 GGAGGAGGAGCCCGACAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 GGAGCGAGCGCGGTTGGCGAGGGTCTGGGCCACACCGGCGAGGACGCGTGGACCGAG 852  
863 TCTGTCTGGTCCCGCCAGGTGCTACT-----GAGAGAAAGATTTGTCTTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTACCTGCCAGACCGCGGAGAGCCACCTCTTTGA 912  
914 AGGAAGGTGTCTACCTGAGTCTCTC-----TGGGTGGTGTGTGTAAACACAGCCAG 970  
913 GGTGTGCTCTCTGGCAGCGGCACCTCCACCCATCCGTGGGCCCGCAGCACACGCGGG 972  
971 CTCACAT-----CTCTGTGTACACCGCCCGCCAAATGCTTTAGCTCAGGCCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCGACACAGTCTCTGGGACACGCTTGTCCCGGTGTAGC 1032  
1028 TGAGCAGAGATTTCTTTACTTCCAGGAGATGCCAAGAGCGTCTAAAACCCCTCAT 1087  
1033 CGAGCAAGCAGCTTCTCTACTCTCTCAGCGA---CAAGGAGCAGCTCGGCCCTCTCT 1089  
1088 CTTACTCAGCAACTCCAGCTAACTTTGACTGGGGCAGAGACTGGTGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCCAGCTGACTGGCGTCCGAGGCTCGTGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGAGCACCTCTGCGAGACACACCGTCTATCGGTG 1207  
1150 TCTGGGTTCCAGGCTCGATGCGAGGACTCCCGCAGGTTGCCCGCTCGCCAGCG 1209  
1208 ATACTGGCAGATCGGCCCTGTTTCCAAAGCTGCTGGTGAACCATGACAGTGGCAATA 1267  
1210 CTACTGGCAATGGGCCCTCTGTTCTGGAGCTGCTTGGGAALCAGCGCAGTGGCCCTA 1269  
1268 TGTGAGACTCTCAGGTCAATTGAGGTTTGAACAGCAAAACCAAGGTGACAGTGC 1327  
1270 GGGGTGCTCTCCTAAGACGACTGCCCCGCTGCGAGTGCAGTCCCGCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGCC 1342  
1330 CTGTGCCGAGAGAGCCCGAGGCTCTGTGGCGGCCCGAGAGGAGGACAGAGCC 1389  
1343 ACCGCACTCTATGATTTGCTCCGCTTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CGTCCGCTGTGTGAGTGTCTCCCGCAGCAGCAGCGCCCTGCGAGGTGTACGGTCTGT 1449  
1403 TCGGCTGTCTGTGCAAGTGTGTCTGTAGTCTCTGGGGTACCAGGACAAATAGAGG 1462  
1450 GCGGCTCTGCTGCGCGGCTGTGCCCCCAGGCTCTGGGGCTCCAGGACAAAGAACG 1509  
1463 CGCTTCTTTAAGAACTTAAAGAGTTTCACTCTGTTGGGAAATACGCAAGTATCACT 1522  
1510 CGCTTCTCAGGAACACCAAGAGTTCACTCTCCTGGGGAAGCATGCCAAGCTCTGCT 1569  
1523 GCAGGAATGATGTGGGAAGATGAAGATAGAGGATTTGCCACTGGCTCGGAGAGCCCGG 1582

1570 GCAGGAGCTGAGCTGGAGATGAGCTGCGGAGCTGCGCTGGCTGCGCAGAGCCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGAGACCGTCTGAGGAGAGAGATCTTGGTACTGTTCT 1642  
1630 GGTGCTGTGTTCGGCGCGCAGAGACCGTCTGCTGAGGAGATCTTGGCCAAAGTTCT 1689  
1643 GTTCTGGCTGATGAGACATACGTGTAGCTGTAGGTCACTTTTATACATACAGA 1702  
1690 GCACTGGCTGATGAGTGTGTAGCTGTGAGCTGTCTAGGTCTTTTATATGTCAGGA 1749  
1703 GAGCACTTCCAGAAAGACAGGCTCTTTCTAGCGTAAAGATGTGTGAGCAAGCTGCA 1762  
1750 GACCACTTCAAAGAAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCACTGAGTCAAGCAACACTTGTAGAGAGTGGCTACGGGAGTGTCTACAAGAGA 1822  
1810 AGCACTTGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGAGCTGTTCGGAAGCAGA 1869  
1823 GGTGAGGCTATCCAGGACACCTGTGCTAGCCATGCCATCTGAGACTTGGCTTTCATCCC 1882  
1870 GGTGAGGAGCATCGGAGAGCCAGGCCCTGCTGAGTCCAGACTCCGCTTTCATCCC 1929  
1883 CAAAGCCAAAGGCTGCGGCCCATTTGTGAACATGAGTTATAGCATGGTACAGAGCTTT 1942  
1930 CAAAGCTGAGCGGCTGCGGCCGATTTGTGAATGGAATACGCTGCGGAGCCAGAGCTT 1989  
1943 GGCAGAAAGAAAGCAGGCCACCATTTCAACAGCGCTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGCACTGTTTCAGCGTCT 2049  
2003 CAACTATAGCGGACAAACATCTCACTTATGGGGTCTTCTGTATCTGGTATGAATGA 2082  
2050 CAACTACGAGCGGCGCGGCCCTCTGCGGCCCTCTGTGTGGGCTTGGACGA 2109  
2063 CATCTACAGACTGCGGGGCTTTGTGCTGCTGCTGCTGCTGAGACACAGACCCAG 2122  
2110 TATCCAGAGGCTTGGCGCACCTTGTGCTGGTGTGGGCCACAGACCCGCGCTGA 2169  
2123 GATGTACTTTGTAAGGCAGATGTGACCGGGGCTTATGATGCCATFCCCGCCAGGTAAGCT 2182  
2170 GCTGTACTTTGTAAGGTGGATGTGACGGGCGGTACGACACCATCCCGCCAGGACAGCT 2229  
2183 GGTGAGGTTGTGTCCATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CAGGAGGTCTCGCCAGCATCATCA---AAGCCAGAACACAGTACTCGGTGCGTGGTA 2286  
2243 TGCACTGGTCCCGAGAGATAGCAAGGCCAAGTCCCAAGTCTCTTTAGGAGACAGGTCCAC 2302  
2287 TGCGTGGTCCAGAAAGCGGCCCATGCGGACAGTCCGCAAGGCTTCAAGAGCCAGTCTC 2346  
2303 CAGCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAAGCATCTGCGAGATTGTA 2382  
2347 TACCTTGAACAGCTCCAGCCGCTACATGCGACAGTTCGTGGCTCACCTGCAATGAGCCAG 2400  
2363 TGCACTGCACTGAGAACTCCGTTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACAGCCCGCTGAGGATGCGCTGTCTATCGAGCAGAGCTCTCCCTGATGAGCCAG 2460  
2423 CAGCAGCTGTTGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCAGCTCTCTCTACGCTTCTATGTCACCAAGCTGTCGCGCATCAGGG 2520  
2483 CAGGTGCTATACAGTGCAGGAGATCCCCCAGGCTCCAGGCTATCCAGCTCTCTG 2542  
2521 CAACTCTAGTCCAGTGCAGGGATCCCGAGGCTCCATCTCTCTCAAGCTGCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAACAGCTGTTGCTGAGGTGCAAGCGGATGCGGT 2602  
2581 CAGCTGTGTACGGGACATGAGAAACAGCTGTTTGGGGATTCGCGCGGACCGGCT 2640  
2603 GCTTTTACGTTTGTGTGATGCTTTCTGTTGTGAGCTCTCTGCTGAGCCAGCAAGAAC 2662



2641 GCTCCTCGGTTTGGTGGATGATTTCTTGTGGTGACACCTACCTACCCAGCGGAAAC 2700  
2663 CTTCTCAGCACCCCTGGTGCATGCGGTTCTAGTATGGGTGATGATTAATACTCCAGAA 2722  
2701 CTTCTCAGGACCCCTGGTCCGAGGTGCTCCCTAGTATGGTGGTGAATTCGGAA 2760  
2723 GACAGTGTGAACTTCCCTGTGGAGCTGTATCCCTGGGTGGTGGAGTCCATACAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTAGAGAGGCGCTGGGTGGCAGGCTTTTGTTCAGAT 2820  
2783 GCTGTCTCACTGCTGTTTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2842  
2821 GCGGCGCCAGCGGCTATTCCCTGGTGGGCGCTGCTGTGGATACCGGACCGTGAAGT 2880  
2843 GTTCTGTGACTACTCAGGTTATCCGAGAGCTCAATTAAGACGAGCCTCACCTTCCAG 2902  
2881 GCAGAGGCACTACTCCAGCTATCCCGGAGCTCCATCAGAGCGAGTCTACCTTCAAC 2940  
2903 TGTCTTCAAGCTGGGAAGACCATGGGGAACAGCTCTCTGCTGCTTGGGTTGAAGTG 2962  
2941 CGGCTTCAAGGCTGGGAGGAACATGGTGGCAACTCTTTGGGCTTGGGCTGAAGTG 3000  
2963 TCAGGCTTATTTCTAGACTTGACGTGACGTGACAGCTCCAGAGTCTGATCAATATA 3022  
3001 TCAGGCTTATTTCTGATTTGAGTGAACAGCTCCAGAGCTGTTGACCAACATCTA 3060  
3023 CAAGATCTTCTCTCAGGCTTACAGGTTCCATGATGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTTCTCTCAGGCTTACAGGTTTACGATGTGCTGAGCTCCCTTTCA 3120  
3083 CCAGGCTTGTAGAGAACTTCACTTCTTGGGCACTATCTCCAGGCAAGCATTCCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCCACTTTTCTGCGGCTCATCTCTGACAGGCTCCCT 3180  
3143 CTGCTATGCTATCTCTGAGGTCAAGATCCAGGATGACATTAAG-----GCTC 3193  
3181 CTGCTATCTCATCTCTGAGGCAAGAACGAGGATGTGCTGGGGCCAGGCGCGC 3240  
3194 TGGCTCTTCTCTCTGAGAGCGCAATGGTCTGTGTACACAGGCTTCTCTGTCAAGCT 3253  
3241 CGGCTCTCTCTCTGAGGCGGTGAGTGGTGTGCCAACCAAGCATTCCTGTCAAGCT 3300  
3254 GGTGTCTCATCTGTCTATCTACAAATGTCTCTGGAGCTCTGAGGACAGCCCAAACT 3313  
3301 GATCGNACCGTGTCTACCTAGTGCCATCTCTGGGTCATCTCAGGACAGCCCAAGCA 3360  
3314 GCTGTGCGGAACTCTCCAGAGCGCAATGACCATCTCTTAAAGTGCAGCTGAGCCAGC 3373  
3361 GCTGAGTGGAACTCTCCGAGGACGAGCTGACTGCTGTGAGGCGCGAGCCACCGGC 3420  
3374 CTTAAGCACAGCTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGACTGA 3454

ULT 7

46821

AAD46821 standard; cDNA; 4015 BP.

AAD46821;

27-JAN-2003 (first entry)

Human telomerase reverse transcriptase (TERT) cDNA.

Carbohydrate antigen; alpha(1,3)galactosyltransferase; alpha1,3GT; TERT;  
transgenic; alpha(1,2)fucosyltransferase; alpha1,2FT; human; enzyme;  
telomerase reverse transcriptase; gene; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 56..3454

PT FT /\*tag= a  
XX XX /product= "Human telomerase reverse transcriptase"  
PN WO200274948-A2.  
XX 26-MAR-2002.  
XX 21-MAR-2002; 2002WO-CA00378.  
XX 21-MAR-2001; 2001US-277811P.  
PR (GERO-) GERON CORP.  
XX Denning C, Clark AJ, Schiff JM;  
XX WPI; 2002-759995/82.  
DR P-PSDB; AAE29226.  
XX Mammalian cells, useful for producing animal tissues with carbohydrate  
PT antigens that are compatible for transplantation into human patients -  
XX Disclosure; Page 33-34; 71pp; English.  
XX The invention relates to animal tissues with carbohydrate antigens that  
CC are compatible for transplantation into human patients. The mammalian  
CC cell is inactivated homogeneously for expression of alpha(1,3)galactosyl-  
CC transferase (alpha1,3GT) gene and comprises a transgene for alpha(1,2)-  
CC fucosyltransferase (alpha1,2FT). It is useful for producing animal tissue  
CC with carbohydrate antigens that are compatible for transplantation into  
CC human patients. The present sequence is human telomerase reverse  
XX transcriptase (TERT) cDNA used in the invention.

SQ Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 24; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

Qy 23 CCAGCGCTTGAGCAACAATGACCGCGCTCTCTGTTGCGCGCGTCTCTGTCTGTGG 82  
Db 40 CCAGCGCAACCCCGGATGCGCGCTCCCGCTCCCGTCCGAGCGGTGCGCTCCCTGTGG 99  
Qy 83 CAGCGGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 142  
Db 100 CAGCGCACTACCGGAGGTGTGCGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 159  
Qy 143 CAGCGCGCTTGTGCAACCCCGGAGCGCGGAGATCTACCGCACTTTGTTTCCCAATGCCT 202  
Db 160 CTGGCGGCTGTGTGACAGCGCGGAGCGCGGCTTTCGGCGGCTGTGTGCGCCAGTGCCT 219  
Qy 203 AGTGTGATGCACTGGGCTCAGAGCTCCAGCTCCAGCTTTCCTTCCAGAGGTGTC 282  
Db 220 GGTGTGCGTGGCTTGGGACGACCGCGCGCGCTTCTCTTCCCGCGCGCTTCCCGAGGTGC 279  
Qy 263 ATCCCTGAAGAGCTGTGGCGGCTGTGTGACAGACTCTGCGAGCGCAACAGAGAAA 322  
Db 280 CTGGCTGAAGAGCTGTGTGGCGCGGAGTGTGACAGCGCGCGCGCGCGAGAA 339  
Qy 323 CGTGTGCGCTTGTGGCTTGTAGCTGTGTTAAAGCGCAGAGGCGCTTCCCATGCGCTT 382  
Db 340 CGTGTGCGCTTGTGGCTTGTGGCTGTGACCGGCGCGCGCGCGCGCGCGCGCTT 399  
Qy 383 CACTAGTAGCGTGTGCTACTTGTCCCAACACTGTTATTGAGACCTGTGTGCTGAGTGG 442  
Db 400 CACCAACAGCTGTGCGAGCTACTTGTCCCAACAGCTGTGACCGAGCGCGCGGAGCGG 459  
Qy 443 TCATGTAGTCTACTTGTGAGCGGAGTGGCGAGCACTGCTGTCTACTGTGTCGACA 502  
Db 460 GCGGTGGGCGCTGTGCTGCGCGCGCTGGGCGAGCACTGTGTGCTTCACTGTGTGACG 519  
Qy 503 CTGTGCTCTTTATCTTCTGTGGCGCGCGCGCTGTGCTGCTACAGGTGTGTGCTCTCCCT 562  
Db 520 CTGCGCGCTCTTTGTGTGGTGGCTTCCAGCTGCGCTTACAGGTGTGCGGCGCGCGCT 579

563 GTACCAAAATTTGTGCACACACGGAATCTGTGGCCCTCTGTGTCGGCTAGTTACAGGCCAC 622  
580 GTACCAAGCTCGGCGCTGCCACTCAGGCCCGGCCGCCACACGCTAGTGGACCCCGAAG 639  
623 CCGACCCGTGGGAGGAAATTCACCTAAGCTTAGTTCTTTACAAACAGATCAAGACAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGGCTGGAACCATAG 672  
683 TCGCCAGAGACCGGAACCCCTGGCTTGCCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 CGTCAGGAGCGCGGGTCCCTCGGCTTGCAGCCCGGGTGGAGAGCGCGGGGG 732  
743 TCTCACCAGTACAAAGTGTGCTTCACTAAGAGGCGCATCTATCTCTCCGAGAGT 802  
733 CAGTGCACGCGGAGTGTGCTTGCCTGAGAGGCCAGCGTGGCGTGCCTCTGAGCC 792  
803 GGAGGAGGACCCACAGCAGTGTGTACAAACCCCATCAGCAAAATCATGGGTCCCAAG 862  
793 GGAGCGGACCGCGTGGGAGGGGTCTGGGCCACCCCGGCGAGGACGCGTGGACCGAG 852  
863 TCCTGCTCGGTCCTCCGAGTGCCTACT-----GCAGAGAAAGATTGTCTCTAA 913  
853 TGACCGTGGTTCGTGTGTGTGTCACCTGCGCAGACCGCGGAGAGCCACCTCTTTGGA 912  
914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTCGGTGTGCTGTATAACACAAAGCCGAG 970  
913 GGGTGGCTCTCTGGCAGCGCCACTCCACCCCATCGTGGGCCGCGCAGCACACGCGGG 972  
971 CTCACAT----CTCTGCTGTACACCGCCGCCAAATGCTTTAGCTCAGGCGCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCCACCACTGCTGGGACACGCTCTGCGCGGTGTACGC 1032  
1028 TGACACAGACATTTCTTTACTTCCAGGGGAGATGCCAAAGAGCGTCTAAACCCCTCATTT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCCAGCTTAATTTGACTGGGCCAGGAGACTGGTGGAGATCATTT 1147  
1090 CTTACTCAGCTCTCTAGAGCCAGCGCTGACTGGCGCTCGGAGGCTCTGTGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGGACCATCTTCAGGACACACCGCTATTCGCGTCG 1207  
1150 TCTGGGTTTCAGGCCCTTGATGTCAGGGACTCCCGCAGGTTGCGCGCTCCGCCAGCG 1209  
1208 ATACTGCAGATCGGCCCTCTGTTCACACAGCTGCTGTGTAAACCAATGCAGAGTGCCAATA 1267  
1210 CTACTGGCAAAATGCGGCCCTGTCTTCTGAGCTGCTTGGGAAACACGCGCAGTGCCTTA 1269  
1268 TGTCAAGTCTCAGGTCAATTGCAGGTTTCGAAACAGCAAAACACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCACCCCAGCAGCGGGT 1329  
1328 CT-----TGAAACACAGGCC 1342  
1330 CTGTGCGCGGAGNAGCCACAGGCTCTGTGGCGGCCCGGAGGAGGACACAGACCC 1389  
1343 ACCGCACTCATGATTTGCTCGCTGACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402  
1390 CCGTGCCTGTGTGAGTGTCTCGGCCAGCACAGCAGCCCTCGCAGGTGTACGGCTCGT 1449  
1403 TCGGGCTGTCTGCAAGTGTGTGCTAGTCTCTGGGTACGAGGACCAATGAGCG 1462  
1450 GCGGCTGCTGCGCGGCTGTGTGCCCCAGGCTCTTGGGGCTCCAGGACCAAGACG 1509  
1463 CCGCTTTTAAGAACTTAAGAGTTCTCTCGTTGGGGAATATCGGAAGCATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAGTTCTCTCCCTGGGGAAGCATGCAAGCTCTCGCT 1569  
1523 GCAGGAACATGTGGAAGATGAATAGAGGATTCGCACTGGCTCCGAGCAGCGCCGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGTTCGGCTGCGGAGGCCAGG 1629

1583 GAAGGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGGAGTCTGGCTAGCTTCCT 1642  
1630 GGTGGCTGTGTTCGGGCCGAGAGCACCGTCTCGTAGAGATCTTGGCCAAAGTTCCT 1689  
1643 GTTCTGGCTGTATGGACACATACGTGTGATAGCTGTAGTTCATTTTACATCACAGA 1702  
1690 GCACTGGCTGTATGAGTGTGTAGCTGTGAGCTGTCTAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCATTTCAGAGAACAGGCTCTTCTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTGTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTGAGTTCAGGCAACACCTTGAGAGAGTGGCTACGGAGCTGTCAACAAGAGGA 1822  
1810 AAGCATTGGAATCAGACAGCACTTGAAGAGGTGCAGCTGCGGAGCTGTGCGAAGCAGA 1869  
1823 GGTAGGCMATCCAGGACACCTGTGCTAGCCATGCCCATCTGCAGACTGCGCTTCATCCC 1882  
1870 GGTACGGCAGCATCGGGAAGCCAGCCCGCTGTGTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTTGCGCCCATTTGTGAACATAGTTTATAGCATGGTACCCAGAGCTTT 1942  
1930 CAAAGCTGACGGCTGCGGCCGATTGTGAACATGACATACGTCTGTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAAGCGAGGCCAGCAATTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGGAGCGTCTCACCTCGAGGTGAAGGCACATGTTTCAGCGTCT 2049  
2003 CAACTATGAGCGGACAAACATCTCTCAGCTTATGGGTCTTCTGCTAGGTATGAATGA 2062  
2050 CAACTACGAGCGGGCGCGCGCCCGCTCTCTGGCGCTCTGTGCTGGGCTTGGAGGA 2109  
2063 CATCTACAGGACTCTGCGGGCCCTTTGTGCTGCGTGTGCGTCTCTGGACACAGACCCAG 2122  
2110 TATCCACAGGCTTGGCGCACCTTCGTGCTGGGTGCGGCCGAGGACCCGCGCTGA 2169  
2123 GATGTAATTTGTAAGGACAGATGTGACCGGGGCCCTATGATGCCATCCCGAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGGATGTGACGGGCGCTACGACACCATCCCGCAGGACAGCT 2229  
2183 GGTGGAGGTGTGTGCAATATGATCAGGCACCTCGGAGAGCAGCTACTGTATCCGCCAGTA 2242  
2230 CACGGAGGTCTCGCCAGCATCATCA---AAGCCAGAACACGTACTGCGTGCCTCGTA 2286  
2243 TGCAGTGTCTCGGAGAGATAGCCAAAGCCAAAGTCCCTTTTAGGAGACAGGTCAAC 2302  
2287 TGGGTGTCTCAGAAAGCGCCCATGGCACGCTCGCAAGGCCCTTCAAGAGCCACGCTCT 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAAGCATCTGCAGGATTCCAGA 2362  
2347 TACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGTGATGAGAACTCCGTTGTTCATCGAGCAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGGATGCGTCTCATCGAGCAGAGCTCTCCCTGAAATGAGGCGAG 2450  
2423 CAGCAGCTGTTTGTGCTTCTTCCTGCACTTCTCGGTACAGTGTGCTGAAGATTGTGTA 2482  
2461 CAGTGGCTCTTTCGACGCTTCTCTACGCTTCATGTGCCACCCAGCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGCAAGTGCAGGGCATCCCCAGGGCTCCAGCCCTATCCACCTGCTCTG 2542  
2521 CAAAGTCTACGCTCAGTGCAGGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAGAACAGCTGTTGCTGAGGTGAGCGGGATGGGT 2602  
2581 CAGCTGTGCTACGCGCATGAGAGAACAGCTGTTGCGGGGATTCGGCGGAGCGGGCT 2640  
2603 GCTTTTACGCTTTTGTGATGCTTTCTGTGTGACGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGTGTGATGATTTCTTGTGTGACACTCACCTCACCCACGCGAAGAC 2700  
2663 CTTCTCTCAGCACCCCTGGTCCATGCGGTTCTGTAGTATGGGTGCAATGATAAACTTGCAGAA 2722



673 CGTCAGGGAGCGGGGCTCCCTCGGCGCTGCGAGCCCGGGTGGAGAGCGCGGGG 732  
743 TCTCACAGTACAAAGTGTGCTTTCAGCTAAGAGGCGAGATGCTATCTGTCGCGAGAGT 802  
733 CAGTGCACCGGAGTCTGCGTGTGCCAAGAGGCGAGCGTGGCTGCGCTGAGCC 792  
803 GAGAGGGGACCCCAAGAGAGTGTGTACAAACCCCATCAGGCAATCATGGGTGCCAAG 862  
793 GAGAGGAGACCGCGGTTGGGAGGGGTCTGGGCGCCACCGCGGAGAGCGCGTGGACGAG 852  
863 TCTGCTCGGTCGCCCGAGTGTCTACT-----GCAGAGAAGATTTGCTCTTAA 913  
853 TGACCGTGGTTCTGTGTGTGTCACTGCCAGACCCGCGAAGAGCCACTCTTTGA 912  
914 AGAAGAGTGTGTGACCTGAGTCTCTC-----TGGGTGCGTGTGCTGTAAACACAGCCAG 970  
913 GGGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGCGCCAGCACACGCGGG 972  
971 CTCACAT---CTCTGTGTACACCCCGCCAAATGCTTTACAGCTCAGGCCATTTAT 1027  
973 CCCCCATCACATCGCGCCACACAGTCTCTGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTATCTCCAGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGAGCAGCTGCGGCCCTCCT 1089  
1088 CCTACTCAGCAACTCCAGCTTAATTGACTGGGCGCAGGACCTGGTGGAGATCATCT 1147  
1090 CCTACTCAGCTCTCTGAGGCCCGAGCTGAGTGGCGCTCGAGGCTCGTGGAGACCATCT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGGACCACTCTGCAAGGACACACCGTCTATCGCGTCG 1207  
1150 TCTGGTTCCAGGCGCTGGATGCCAGGACTCCCGCAGAGTTGCCCGCGCTGCCCCAGCG 1209  
1208 ATACTGCAATGCGGCGCTGTTCMAAGCTGCTGTGTGACCACTGAGAGTGCACATA 1267  
1210 CTACTGCAATGCGGCGCTGTCTTCTGAGCTGCTTTGGGAAACCAAGCAGTGGCCCTA 1269  
1268 TGTCAAGCTCTCAGGTCAATTGCAAGTTTCAAGCAGAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGAGCACTGCCCGCTGCGAGCTGCGGTGACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGGCC 1342  
1330 CTGTGCCCGGAGAACCCAGGCTCTGTGGGCGCCCGCGAGGAGGACACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402  
1390 CCGTGCCTGTGAGCTGTCTGCCAGCAAGAGCCCTGSCAGTGTACGGCTTGT 1449  
1403 TCGGCGCTGTCTGCAAGTGTGTGTGTGTCTGTGCTGTGGGTACAGGCACAATGAGCG 1462  
1450 CGGCGCTGTCTGCGCGGTGTGTGCGCCAGGCTCTGGGGTCCAGGCAACAAGACG 1509  
1463 CGGCTCTTTAAGAACTTAAGAGTCTATCTGTTGGGAATACGGAAGTCACT 1522  
1510 CGGTTCTCTAGGAACCAAGAAAGTTCTCTCCCTGGGAGACATGCCAAGCTCTGCT 1569  
1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATGTCACTGTGCTCCGAGCAGCGCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGAGTGTGCTGTGCGTGGAGAGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGAGAGCAAGCTGTGAGGAGAGGATCCTGGCTACGTTCT 1642  
1630 GGTGTGCTGTCTCCGCGCAGAGCAAGCTGTGCGTGGAGGATCCTGGCCAAAGTTCT 1689  
1643 GTCTGGCTGATGACACATACGTGGGTACAGCTGCTTAGTCTCTTTTATACACAGA 1702  
1690 GCACTGGCTGATGAGTGTGACGTGCTGAGCTGCTCAGTCTTTCTTTTATGTACGGA 1749  
1703 GAGCACATTCAGAGAACAGGCTCTTCTTCTACCGTAAGAGTGTGTGAGCAAGCTGCA 1762

1750 GACCAGCTTCAAANGAACAGGCTCTTTTCTACCGGAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATGTGAGTCAAGCAACACCTTTGAGAGAGTGGGCTACGGAGAGTGTCAACAGAGGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTTGAAGAGGCTGAGTGGGAGCTGTCCGAAGCAGA 1869  
1823 GGTCAAGCATCACAGGACACCTGGCTAGCCATCCCATCTGCAAGACTGGCTTCATCCC 1882  
1870 GTTCAGCGCAGCATCGGAGAGCGGCGCCCTGCTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGCGCTGCGGCCCATTTGTGAACATGAGTATATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTGACGGGCTGCGGCCGATTTGTGAACATGAGTCTACCTGCTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGAGGCGCCAGCAATTCACCCAGCGTCTCAAGACTCTCTTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGCTCTCACTCAGAGGTGAGGCACTGTTTCAGCGTGTCT 2049  
2003 CAACTATGAGCGGACAAACACATCTTCACTTATGGGTCCTTCTGTATGGGTATGAATGA 2062  
2050 CAACTACGAGCGGCGCGCGCGCTCTGCGGCGCTCTGCTGGGCGCTGTGCTGGGCGTGGACGA 2109  
2063 CATCTACAGGACCTGGCGGCGCTTTGTGCTGCTGCTGCTGCTGAGCCAGACACCCAG 2122  
2110 TATCCACAGGGCTGGCGACCTTGTGCTGCTGCTGCTGCGGCCCGCAGGACCGCGCGCTGA 2169  
2123 GATGTACTTTGTTAAGCGAGATGTGACCGGGCGCTTATGATGCCATCCCGCAGGGTAAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGTGATGTGACGCGCGCTAGACACCATCCCGCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGAGGTCTATCGCCAGCATCATCA---AAGCCGAGAACACGTAAGTCTGCGTGGGTA 2286  
2243 TGCAGTGTGCGGAGAGATGACAAAGGCCAAAGTCCACAAGTCCCTTTAGGAGACAGGTCAC 2302  
2287 TGCCTGTGTCAGAGGCGCGCCCTAGGCAAGTCCGCAAGGCTTCAAGAGCCAGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAAGCATCTGCAAGATTGAGA 2362  
2347 TACCTTGAACAGACCTCCAGCCGTACATGCAAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCACATGAGGACCTCGTGTGATGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACACGCCGCTGAGGAGATGCGCTGCTATGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTGTGAATTTCTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482  
2461 CAGTGGCTCTTTCGACGCTCTTCTTACGCTTCTATGTCACACCGCGCTGCGCATCAGGGG 2520  
2483 CAGTGTCTATACGAGTGCAGGCGATCCCGCGGCTGCGAGCTATCCAGCTGCTCTG 2542  
2521 CAAAGTCTACGTCAGTGCAGGAGATCCCGCAGGCTGCTCTCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTCGAGAGACATGAGAACAAAGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGGACATGAGAACAAAGCTGTTGCGGGATTCGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGAGCGCTGCTGAGCAGCTGCTGAGCAGCAAGAAAC 2662  
2641 GCTCTGCTGTTGGTGGATGATTTCTTGTGTTGAGTCACTTCACTTCACTTCACTTCACTTCACT 2700  
2663 CTCTCTCAGCAGCCCTGCTCCATGGGCTTCTGAGTATGGGTGCTGATGAATACTTCAGAA 2722  
2701 CTCTCTCAGCAGCCCTGCTGCGAGGCTGCTCTGAGTATGGCTGCGTGGTGAATTCGCGAA 2760  
2723 GACAGTGTGAATTTCCCTGTGAGGCTGCTGTAACCTGGTGGTGGAGCTTCCATACAGCT 2782  
2761 GACAGTGTGAATTTCCCTGTGAGGCTGCTGTAACCTGGTGGAGCTTTCCTTTCAGAT 2820  
2783 GCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCAGGCGCTTATTCCTTGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880

2843 GTTCTGTGACTACTAGGTATGCCCCAGACCTCAATTAAGACGAGCTCAGCTTCCAGAG 2902  
2881 GCAGAGCGACTACTCAGGTATGCCCCGACCTCCATCAGAGCGAGTCTCAGCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAACACCACTCGGAACAAGCTCTCTGCGGTCTTGGCGTTGAAGTG 2962  
2941 CGCTTCAAGCTGGGAACAACATGGCTCGCAACTCTTTGGGTCTTGGCGTTGAAGTG 3000  
2963 TCAGGTCTATTCTAGACTTGGAGTGAAGCCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCAGAGCTGTTTCTGGATTGCGAGTGAACAGCCTCCAGACGCTGTGCACCAATCTA 3060  
3023 CAAGATCTTCTGCTTCCAGGCTACAGGTTCCATCATGTGATTCAGTCTTCCCTTTGA 3082  
3061 CAGATCTCTCTGCTGCGAGCGTACAGGTTTACGCATGTGTGCTGAGCTCCCATTTCA 3120  
3083 CCAGCTGTTAGGAAGAACTCACATCTTTCTGGGATCATCTCCAGCCAAAGCATCTTG 3142  
3121 TCAGCAAGTTTGAAGACCCACATTTTCTGCGCGTCTATCTGACACGGCTCCCT 3180  
3143 CTGCTATGCTATCTCAAGTCAAGATCCAGATGACACTAAAG-----GCTC 3193  
3181 CTGCTACTCTCACTCTGAAAGCCAGAACGACAGGATGTGCTGGGGCCAGGCGCGC 3240  
3194 TGCTCTCTTCTCTGAAAGCCGACATTTGCTCTGACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTCTGCGAGCGGTGAGTGGTGTGCGCAAGCATTTCTGCTCAAGCT 3300  
3254 GGTGCTCATCTGCTCATCTACAAATGCTCTGAGACTCTGAGGACAGCCAAAAC 3313  
3301 GACTGACACCGGTGCTACCTAGTGGCTCTGCGGCTCACTCAGGACAGCCAGCGCA 3360  
3314 GCTGTGCGGAAGCTCCAGAGCGGACAAATGACCATCTTAAAGCTGAGTGACCCAGC 3373  
3361 GCTGAGTGGAGCTCCCGGAGGAGCGTGAAGTCTGCTGAGGCGCGCAGCAACCCGCG 3420  
3374 CTTAAGCAGAGCTTTTCAGACCAATTTTGGACTAA 3407  
3421 ACTGCTCTCAGACTTCAAGACCATCTGGAAGTGA 3454

ILT 9  
22474  
ABZ22474 standard; cDNA; 4015 BP.

ABZ22474;

25-MAR-2003 (first entry)

Human telomerase reverse transcriptase encoding cDNA SEQ ID NO:1.

Human; telomerase reverse transcriptase; enzyme; hTERT; chromosome 5; vulnary; antitumor; epithelial cell migration promoter; wound; epithelialization; skin wound; lesion; burn; surgical incision; ulcer; epithelial cell; keratinocyte; epidermal; mucosal; gene; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 56..3454  
/\*tag= a  
/product= "human telomerase reverse transcriptase"

WO200291999-A2.

21-NOV-2002.

09-MAY-2002; 2002WO-US14867.

09-MAY-2001; 2001US-289903P.

(GRO-) GERON CORP.

XX  
PI Jiang X, Chiu C, Harley CB;  
XX WPI; 2003-120591/11.  
DR P-PSDB; ABP56676.  
XX  
XX Composition for treating wounds and enhancing epithelialization of a skin  
PT surface, comprises vector encoding telomerase reverse transcriptase or  
PT telomerase epithelial cells on a microparticle or a matrix  
XX  
XX Disclosure; Page 31-32; 69pp; English.  
XX  
XX The present invention describes a pharmaceutical composition (I) comprising a vector encoding telomerase reverse transcriptase (TERT) in an excipient or device, or comprises telomerase epithelial cells on a microparticle or a matrix suitable for topical administration or administration to a wound site. (II) has vulnary and antitumor activities and can be used to promote epithelial cell migration. (I) is useful for treating a wound and enhancing epithelialization of a skin surface. The wound is especially skin wound including acute lesion such as traumatic lesion, burn, or surgical incision, chronic lesion such as chronic venous ulcer, diabetic ulcer or compression ulcer and the wound is further monitored for closure. The telomerase activity or TERT expression is increased in epithelial cells at the site of treatment and also in fibroblasts or endothelial cells at the site of treatment. The epithelial cells are especially keratinocytes. A polynucleotide encoding TERT is useful for the preparation of a medicament for treatment of a wound or an epithelial surface in a human or animal. An epithelial cell with increased telomerase activity or increased expression of TERT is useful for preparation of a medicament for the treatment of a wound in a human or animal. (I) is also useful for treating wounds of other epidermal surfaces including mucosal surfaces such as bronchus, mouth, nose, oesophagus, stomach, or intestine. The present sequence encodes human TERT (hTERT), which is given in the exemplification of the present invention. hTERT is located to chromosome 5.  
XX  
XX Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 25; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCCGCCCTTGAGCAATGACCCCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGTCGCG 82  
Db 40 CCCGCCACCCCGGATGCGCGCGCTCCCGGTGCGGAGCGCTCCCTGCTGCG 99  
QY 83 CAGCCGATACCGGAGGTGTGCGCGTGGCAACCTTTGTGCGCGCTGCGGCGCCGAGG 142  
Db 100 CAGCCACTACCGGAGGTGTGCGCGCTGGCCACCTTCTGCGCGCTGCGGCGCCAGG 159  
QY 143 CAGCGCGCTTGTGCAACCCCGGAGACCCGAGATCTACCGCACTTTGGTTGCCCAATGCT 202  
Db 160 CTGGCGCTGTGTGACGCGCGGAGACCCCGCGCTTCCGCGCTGTGTGCGCCAGTGC 219  
QY 203 AGTGTGATGACATGCGGCTCAGAGCTTCCAGCTTCCAGCTTTCCTTCCAGAGTGT 262  
Db 220 GGTGTGCTGCTTGGGACGCAAGCGCGCGCCCGCGCGCTCTCTTCCGCGCGAGTGT 279  
QY 263 ATCCCTGAAGAGCTGTGCGCAGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTGAAGAGCTGTGTGCGCGCTGCTCAGAGGCTGTGCGAGCGCGCGCGAGAA 339  
QY 323 CCGTGTGCTTTTGGCTTTGAGCTGCTTAAAGAGCCAGAGCGCGGCTCCCATGGCTT 382  
Db 340 CCGTGTGCTTGTGCGCTTGTGCGCTGCTGAGCGCGCGCGCGCGCGCGCGCTT 399  
QY 383 CACTAGTAGCTGTGCTAGCTACTTGTGCGCGAGCTGTATTGAGACCTGCTGCTGTCAG 442  
Db 400 CACCAACGCTGTGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 459  
QY 443 TGCATGATGCTACTGTTGAGCGCGAGTGGGCGAGCGCTGCTGCTGCTGCTGCTG 502  
Db 460 GCGGTGGGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 519

503 CTGTGCTCTTTATCTTCTGTGTGCCCCCAGCTGTGCTTACAGGTGTGTGGTCTCCCT 562  
520 CTGGCGCTCTTTGTGTGTGGTCTCCAGCTCGCTTACAGGTGTGGCGCCCGCT 579  
563 GTACCAAAATTTGTCCACCAAGGATATCTGGCCCTCTGTGTCCGCTAGTTTACAGGCCAC 622  
580 GTAACAGCTCGGCGCTGCCACTCAGGCCCGCCCCCGCCACACAGCTAGTGTGACCCGGA 639  
623 CCGACCCGTGGGCGAGGAATTTCACTAACCTTAGGTCTTTACAACAGATCAAGAGCGTAG 682  
640 CGCTCTGGGATGC-----GACCGGCCCTGGAACCATAG 672  
683 TCGCCAGGAGCACCGAAACCCCTGGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 CGTCAGGAGGCGCGGGTCCCTTCGGGCTGCCAGCCCCGGGTGCGAGAGGCGCGGG 732  
743 TCTCACAGTACAAGTGTGCTTCAAGTAAAGAGGCGAGATGTATCTGTCCCGAGAT 802  
733 CAGTGCAGCGGAAGTCTGCGGTTGCCAAGAGGCCCGAGCGTGGCGCTGCCCTGAGCC 792  
803 GGAGGAGGACCCACAGGAGGTGCTTACCAACCCATCAGGCAATCATGGTGCACAG 862  
793 GGAGCGAGCGCCCTTGGGCGAGGGTCTTGGGCCACCCGGGCGAGCGCGTGGACCGAG 852  
863 TCCTGCTCGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTTGTCTTTCTAA 913  
853 TGACCGTGGTTTCTGTGTGTGTCACTGCGAGACCGCGGAAGAGCCACCTCTTTTGA 912  
914 AGGAAAGGTGTCACTGATGTCTC---TGGGTGGTGTGTGTAAACAAGCCAG 970  
913 GGGTGGCTCTCTGGCAGCGGCCACTCCACCCATCGTGGGCGCGCAGCACACGCGGG 972  
971 CTCACAT---CTGTGTGTCAACACCGCCGCAAAATGCCCTTTCAGCTCAGGCCATTTAT 1027  
973 CCCCCATCCACATCGGGGCCACACAGTCCCTGGGACACGCTTGTCCCCGGGTACGC 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAAAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGTGGCGCCCTCT 1089  
1088 CTTACTCAGCACTCCAGCCTACTTGTACTTGGGCCAGGAGACTGTGTGAGATCATCT 1147  
1090 CTTACTCAGCTCTCTGAGGCCCGCTCTGATGGGCTCGGAGGCTGTGTGAGACCATCT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGGACCACTCTGCGAGACACACGCTTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCCCTGGATGCGCAGGCACTCCCGCAGGTTGCCCGGCTGCCCGCAG 1209  
1208 ATACTGGCAGATGGGCCCTGTTCACAGCTGCTGTGAACCATGCGAGGTGCCAATA 1267  
1210 CTACTGGCAATGGGCCCTGTTCCTGAGCTGCTTTGGGAACACCGCGAGTGGCCCTTA 1269  
1268 TGTACAGCTCTCAGGTCACTTGCAGGTTTCGAACAGCAAAACCAAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACGCACTGCCGCTCGAGCTGGGTCACCCAGCAGCGGCTGT 1329  
1328 CT-----TGAAACACAGCC 1342  
1330 CTGTGCCCGGAGAACCCAGGCTCTGTGGGGCCCGCGAGGAGGAGACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCCCTGCAAGAGTCCCTGGCAGGTATATGGTTTTCT 1402  
1390 CGGTGCTGTGTGAGTGTCTCCGCGAGACAGAGCCCTTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGCCCTGTCTGTGAAGGTGTGTGTGTGTCTGTGGGGTACAGGCACAATGAGCG 1462  
1450 GGGGCCCTGCTGGCGCGCTGTGTGCCCGCCAGGCTCTGGGGCTCCAGGCAACGAACG 1509  
1463 CGCTTCTTTAAGAACTTAAAGAAAGTTTCACTCTGTGGGGAATATCGGAAGCTATCACT 1522  
1510 CGCTTCTCAGGAACACCAAGAAAGTTTCACTCTGTGGGGAAGCATGCCAGCTCTCGCT 1569

1523 GCAGGAATGTATGTGGAGATGAAGTAGAGATTTGCCACTGGCTCCGCGAGCCCGGG 1582  
1570 GCAGGAGCTGAGGTGAAGATGAGCGTGGGAGCTTGGCTTGGCTGCGCAGAGCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGACCGCTCTGAGGAGAGGATCTCTGGTACGTTCT 1642  
1630 GGTGGCTGTCTCGGCGCAGAGACCGCTCTGCGTGGAGAGATCTCTGGCCAAGTTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGTACAGCTGTAGGTCACTTCTTTTACATCACAGA 1702  
1690 GCACTGGCTGATGAGTGTGTAGTGTGTCTCAGGTCTTTCTTTTATGTCAACGA 1749  
1703 GAGCATAATCCAGAAACAGGCTCTTCTTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTTTCAAAGAACAGGCTCTTTTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTTCAGGCAACACCTTGTAGAGAGTGGGTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCACAGACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGAAGCCAGGCCCGCTGTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTCTCGGCCCATTTGTGAACATGATGATATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTGTACGGCTCGGCCGATTTGTAAACATGACTACGTCGTGGGAGCCAGAACGTT 1989  
1943 GGCAGAGAGGAGCAGGCCAGCATTTACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCCAGAGAAAGAGGCCGAGCGTCTCACTCGAGGTGAAGGACCTGTTTCAGCGTCT 2049  
2003 CAACTATGAGCGGACAAAAATCTCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACTAGAGCGGCGCGCGCCCGCCCTCTCTGGCGCTCTGTGTGGGCTGGAGCA 2109  
2063 CATCTACAGACTGCGCGGCTTTGTGTGGTGTGGTGTCTGGACCGACAGACCCAG 2122  
2110 TATCCACAGGGCTGCGCGACCTTGTGTGTGGTGTGGGCGCCAGGACCCCGCCCTGA 2169  
2123 GATGTACTTTTAAAGCAGATGTACCGGGGCTTATGATGCCATCCCCCAGGGTAAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGTGATGAGCGGCGCTACGACACCAATCCCCCAGGACAGCT 2229  
2183 GGTGAGGTTGTTGCCAATATGATGAGGACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CAGGAGGTCAATCGCAGCATCATCA---AACCCAGAACAGTACTCGTGGCTCGGTA 2286  
2243 TGCAATGTTCCGAGAGATAGCAAGGCCAAGTCCCAAGTCTCTTTAGGAGACAGGTCAAC 2302  
2287 TGCGTGTCCAGAAAGCGCCCATGGGCACTCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CAGCTCTGTGACTCCAGCATACATGGGCGAGTTCTTAAAGTCTGTGAGGATTCAGA 2362  
2347 TACCTTGACAGACTCCAGCCGCTACATGCGACAGTTCGTGGGCTCACTGSCAG-----GA 2400  
2363 TGCCAGTGCATCAGGAACTCCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCGCTGAGGATGCGCTGTCTATCAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGTACTTCTTCTGCACTTCTGCTGCTCAAGTGTCTGTAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTCTGAGCTCTTCTAGCTTCTATGTGCCACCGCCGTGGCATCAGGGG 2520  
2483 CAGGTGTATACAGTGTGCGAGGCTATCCCGAGGCTCCAGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTGTCTGCTGCGAGGGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTGTGTGTTTCGAGACATGAGAAACAGTGTGTTGCTGAGGTGCGAGGATTTGGT 2602  
2581 CAGCTGTGTCTACGGCGACATGAGAGAACAGCTGTTTGGCGGATTCGCGGAGCGGCT 2640  
2603 GCTTTACAGTTTGTGTATGATCTTCTGTGTGTGAGCGCTCTCACTTGGACCAAGCAAAAC 2662





503 CTGTGCTCTTTATCTTTCTGTGCCCCCAGCTGTGCTACAGGTGTGTGGTCTCCCT 562  
520 CTGGCCCTCTTTGTGTGTGTGCTCCAGCTCGCCTTACAGGTGTGTGGCCCTCCCT 579  
563 GTACCAAAATTTGTGCCACCAAGATATCTGTGGCCCTCTGTGCTAGTTTACAGCCAC 622  
580 GTACCAAGCTCGGCGCTGCACCTCAGGCCCGCCCGCCACACGCTAGTGGACCCGAAG 639  
623 CCGACCGTGGGAGGAAATTTCACTAACCCTAGGTCTTACAACAGATCAAGAGCAGTAG 682  
640 CGGTCTGGGATGC-----GAAACCGGCTTGAACCATAG 672  
683 TCGCCAGGAAGACCGAAACCCCTGCTTGCATCTCAGAGGTACAAGAGAGCATCTGAG 742  
673 CGTCAGGAGGCGCGGGTCCCTCTGGGCTTGCAGCCCGGGTGGAGAGGCGCGGGGG 732  
743 TCTCACAGTACAGTGTGCTTCACTAAGAGGCCAGATGCTATCTCTTCCAGAGT 802  
733 CAGTGCAGCCGAAGTCTGCGGTGCCCAGAGGCCAGGCGTGGCGTGCCTCTGAGCC 792  
803 GGAGGAGGACCCACAGGAGGTGTACCAACCCCATCAGCAAAATCATGGGTGCCAAG 862  
793 GGAGCGGACCCCGTGGGACAGGGTCTTGGGCCACCCCGGCGAGGACCGTGGACCGAG 852  
863 TCCTGTCTGCTCCCGAGTGTCTACT-----GCAGAGAAAGATTTGTCTTTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCACTGCCAGACCCGCGGAAGACCACTCTTTGGA 912  
914 AGAAAGGTCTGACCTGAGTCTCTC---TGGGTGCTGTGCTTAACACAGACCCAG 970  
913 GGTGTGCTCTTGGACCGGCCACTTCCACCCATCCGTGGGCCCGCAGCACCGCGGG 972  
971 CTCACAT---CTCTGTGTACACCCCGCCAAATGCTTTTACGTACAGGCCATTTAT 1027  
973 CCCCCATCCACATCGCGCCACCACTCCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTTACTCAGGGAGATGGCAGAGCGTCTAAACCCCTCATTT 1087  
1033 CGAGACCAAGCATTTCTCTACTCTCTCAGGCGA---CAAGGAGAGCTGCGGCCCTCTT 1089  
1088 CTTACTCAGCAACCTCCAGGCTAACTTGACTGGGCGCAGGAGCTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTTGAGGCCAGCTGACTGGCTCGAGGCTCGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGACCACTCTGAGGACACACCGTCTATCGCGTGC 1207  
1150 TCTGGGTTCCAGGCTTGAATGCCAGGAGCTCCCGCAGGTTGCCCGCTTCCCGCAGCG 1209  
1208 ATACTGACAGATGGGCCCTGTGTTCCAGAGCTGTGGTGAACCATGACAGTGGCAATA 1267  
1210 CTACTGCAATGCGGCCCTGTTTCTGAGCTGTCTTGGAACTCAGCGAGTGCCTTA 1269  
1268 TGTCAAGTCTCTCAGTFCATATTGCAAGTTTGAACAGAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGACTGCGCGCTGCGAGTGGGTGCGGTCAACCCAGACCGGTGT 1329  
1328 CT-----TGAAACACAGGCC 1342  
1330 CTGTGCCGGAGAACCCAGGCTCTGTGGGCGCCCGCAGGAGGAGGACACAGACCC 1389  
1343 ACCGCACTCATGATTTGTCTCGGCTTGACAGCAGTCCCTGCGAGGTATATGTTTTCT 1402  
1390 CCGTCCGCTGTGACGTCTCCGCCAGCACAGACGCCCTTGGCAGGTGTACGGCTTCGT 1449  
1403 TCGGGCTGTCTTGAAAGTGTGTGCTGTAGTCTCTGGGTGACAGGCACAATGAGCG 1462  
1450 GCGGGCTGTCTGCGCGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509  
1463 CCGCTCTTTAAGAACTTAAAGAGTTTCACTCTGTTGGGAATACGGAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACCAAGAAAGTTTCACTCTCTGCGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGGAATGATGTGGAAGATGAAGATGAGGATATGCCACTGGCTCCGACGACGCCGG 1582

1570 GCAGGAGCTGACGCTGGAGATGAGCTGCGGAGCTGCGCTTGGCTGGCAGGACCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGACACGCTCTGAGGAGAGATCTCTGGCTAGCTTCCT 1642  
1630 GGTGGCTGTGTTCGGGCGGACAGACACGCTCTGCTGAGAGATCTCTGGCCAAAGTTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGTACAGTCTGTTAGGTCACTTTTACATCAGAGA 1702  
1690 GACTGGCTGATGAGTGTGCTGCTGAGCTGCTCAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCATTTCAGAGAAACAGGCTCTTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTTCAAGAAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATGGAGTGCAGCAACCTTGAAGAGTGGCTGAGAGGCTGTGCGGAGCTGTGGAAGCAGA 1822  
1810 AAGCATGGAATCAACAGACACTTGAAGAGGTGAGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCACAGGACACCTGGCTAGCCATGCCATCTGCAGACTTGGCTTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGAAAGCCAGGCCCGCTGCTGAGTCCAGACTCCGCTTTCATCCC 1929  
1883 CAAAGCCAAACGCTTGGGCGCAATTGTGAACATGAGTTATAGCTGGGTACAGAGCTTT 1942  
1930 CAAAGCTGACGGCTTGGGCGCAATTGTGAACATGAGTACGCTGCTGGAGCCAGAACGTT 1989  
1943 GGGCAAGAAAGAGCAGGCCAGCATTTTCAACAGGCTCTCAAGACTCTTTCAGCATGCT 2002  
1990 CGCAGAGAAAGAGGCGGAGCTCTCACTCGAGGTGAAGCACTGTTTCAGCGTGT 2049  
2003 CAACTATAGCGGACAAACATCTCACTTATGGGTCTTCTGTATCTGGGTATGAATGA 2062  
2050 CAACTACGAGCGGCGCGCGCTCTGCGGCTCTCTGCTGGCTTGGCTTGGAGCA 2109  
2063 CATCTACAGGACTGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCACAGGCTTGGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATGTACTTTGTTAAGGAGATGTGACCGGGGCTTATGATGCCATCCCCAGGGTAAAGCT 2182  
2170 GCTGTACTTTGTAAGGTGATGTGACGGGCGCTGACGACCATCCCCAGGACAGGCT 2229  
2183 GGTGGAGTGTGTCATATGATCAGGACATCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CAGGAGGTCTATCGCAGCATCA---AAGCCAGAACACGACTGCGTGGCTCGGTA 2286  
2243 TGCAAGTGTCCGAGAGATAGCCAGGCAAGTCCCAAGTCTTTAGGAGACAGGTCA 2302  
2287 TSCGCTGTCCAGAGGCGCCCATGCGCAGCTCGGAGGCTTCAAGAGCCAGCTTC 2346  
2303 CACCTCTCTGACCTTCCAGGCAATACATGGGCGAGTCTCTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGCTACATGCGACAGTTCTGGCTCACCTGCGAG-----GA 2400  
2363 TGCAGTGCATGAGGAACTCCGTTGCTCATGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGAGTGCCTGCTCATGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGAATCTTCTGCTCACTTCTCGCTACAGTGTCTGTAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTCGACGCTCTTCTACGCTTCTATGTCACACCGCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATCGCAGTCCAGGAGCATCCCCAGGCTCCAGCTATCCACCTGCTCTG 2542  
2521 CAACTCTACGCTCCAGTGCAGGAGTCCCGAGGCTCCATCTCTCTCAGCTGCTCTG 2580  
2543 CAGTCTGCTTTTCGAGAGCATGAGAAACAGTGTGCTGCTGAGGTGAGCGGGATGGGTT 2602  
2581 CAGCTGTCTACGCGACATGAGAAACAGCTGTTTTCGGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGTATGACTTCTGTTGGTGAAGCTTCTGTTGAGCGCTCAGTTGAGCAAGCAAAAC 2662





2603 GCTTTTACGTTTTTGTGATGACTTTTCTGTGTGTGACGCTCAGCTTGACCAAGCAAAAAC 2662  
2644 GCTCCTGCGTTTTGGTGGATGATTTCTTTGTGTGACACCTCACCTCACCACGCGAAAAAC 2703  
2663 CTTTCCCTCAGACACCTCGTCCATGGCGGTTCTCTGAGTATGGGTGCATGATAAATTTGCAGAA 2722  
2704 CTTTCCCTCAGACACCTCGTCCGAGGTGTCCCTGAGTATGGCTGCGTGTGAACTTTCGGAA 2763  
2723 GACAGTGGTGAATCTCCCTGTGAGAGCTTGTAACCTGGGTGGTGCAGCTCCATACAGCT 2782  
2764 GACAGTGGTGAATCTCCCTGTGAGAGAGAGAGGCCCTGGGTGGCACGGCTTTTGTTCAGAT 2823  
2783 GCTGTCTCAGTGCCTGTTTTCCCTGGTGTGGCTTGGCTGACACTCAGACTTTTGAGGT 2842  
2824 GCGGCCCAACGGCTTATTTCCCTGGTGGCGGCTGCTGCTGGATACCCGGACCTTGAGGT 2883  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAAGACGAGCCTCACTTCCAGAG 2902  
2884 GCAGAGGACTACTCCAGCTATGCCGGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2943  
2903 TGTCTTCAAGAGCTGGGAGACATCGGAAACAGAGCTCCTGTGGCTTGGGTGAGTG 2962  
2944 CGGCTTCAAGGCTGGGAGGAACATGGTCCGAAACTCTTTGGGGTCTTGGGCTGAGTG 3003  
2963 TCAGGCTCTATTTCTAGACTTCGAGCTGAACAGCCTCCAGACAGCTGCTCATCAATATATA 3022  
3004 TCAGAGCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCCAACATCTA 3063  
3023 CAAGATCTTCTGTCTCAGGCTTACAGGTTTCATGATGTGTGATTCAGCTTCCCTTTGA 3082  
3064 CAAGATCTCTGTGTGAGGCGTACAGGTTTCAACGATGTGTGTGTGTCAGCTTCCCATTTCA 3123  
3083 CCAGGCTGTATAGGAAGAACTCATCTTTCTGGGCTCATCTCCAGCCAAGCATCCTG 3142  
3124 TCAGCAAGTTTGGAGAAACCCACATTTTCTGGCGTCTCTGTGACAGGCTCCCT 3183  
3143 CTGTATGCTATCTGAAGGTCGAAGATCCAGAAATGACATAAG-----GCCTC 3193  
3184 CTGTACTCCATCTGTAAGCCCAAGAACCGAGGATGTGCTGGGGGCAAGGGCGCGCG 3243  
3194 TGGCTCTCTTCTCCTCAAGCCGACATTTGGCTCTGCTACAGGCTCTCTGCTCAAGCT 3253  
3244 CGGCCCTCTGCCCTCCAGGCGGTGCAGTGGCTGTGTGCCAACGAATTCCTGCTCAAGCT 3303  
3254 GGCTGCTCATTTCTGCATCTACAAATGTCTCTGGGAOCTCTGAGGAAGGCCAAAAACT 3313  
3304 GACTCGACACCGTGTCACTACGTGCCACTCTCTGGGCTCACTCAGGACAGGCCAGACGA 3363  
3314 GCTGTGCGGGAAGCTCCAGAGGCGCAATGACCATCTTTAAAGCTGCAGCTGACCCAGC 3373  
3364 GCTGAGTCGGAAGCTCCGGGAGCGAGCTGACTGCTCTTGGAGGCGGACGACACCCGGC 3423  
3374 CCTAAGCAACAGACTTTTTCAGACCAATTTTGGACTAA 3407  
3424 ACTGCCCTCAGACTTCAAGACCACTCTGGACTGA 3457

ULT 12  
89424

AAX89424 standard: cDNA: 4027 BP

**AXX89424;**

22-SEP-1999 (first entry)

Human EST2 coding sequence.

ESR2; proliferative capacity; cellular proliferation; decubitus ulcer; telomerase-activating therapeutic agent; cell life-span extension; venous disease; venous stasis ulcer; excessive pressure; arterial ulcer; tissue regeneration enhancer; atherosclerosis; therapy; ss

OS	Homo sapiens.
XX	
XX	WO9935243-A2.
PN	
XX	
PD	15-JUL-1999.
XX	
PF	12-JAN-1999; 99WO-US00682.
XX	
PR	21-APR-1998; 98US-0063557.
XX	
PR	12-JAN-1998; 98US-0071220.
PR	13-JAN-1998; 98US-0071455.
XX	
XX	(COLD-) COLD SPRING HARBOR LAB.
XX	
XX	Beach DH, Hannon GJ, Wang J;
PI	
XX	
DR	WPI: 1999-444196/37.
DR	F-PSDB; AAY28401.
XX	
XX	Increasing proliferative capacity of cells useful for promoting
PT	wound healing
PT	
XX	
PS	Claim 4; Page 65-70; 73pp; English.
XX	
CC	This sequence encodes the human ES2 protein, and can be used in the
CC	method of the invention. The method is for increasing the proliferative
CC	capacity of cells, and comprises contacting the cell with a
CC	telomerase-activating therapeutic agent (TARA). The method can be used
CC	for extending the life-span of cells, e.g. by increasing the number of
CC	mitotic divisions. They can be used for e.g. the extension of skin or
CC	other epithelial cell cultures or grafts, the expansion of mesenchymal
CC	cell cultures or grafts, and the expansion of chondrocyte or osteocyte
CC	cultures or grafts. They can be applied to e.g. neuronal, haematopoietic,
CC	epithelial, pancreatic, hepatic, chondrocytic and osteocytic stem and
CC	progenitor cells in vivo, in vitro or ex vivo protocols. The methods
CC	can be used for promoting the healing of wounds resulting from
CC	e.g. surgery, burns, inflammation or irritation of ulcers resulting from
CC	e.g. venous disease (venous stasis ulcers), excessive pressure (decubitus
CC	ulcers) or arterial ulcers. They can also be used to enhance tissue
CC	regeneration processes, e.g. of the skin, hair and/or fingernails. They
CC	can also be used for treating age-related conditions, e.g. atrophy of the
CC	skin through loss of extracellular matrix homeostasis, in dermal
CC	fibroblasts, age-related macular degeneration caused by accumulation of
CC	lipoic acid and downregulation of a neuronal survival factor in retinal
CC	pigmented epithelial (RPE) cells, and atherosclerosis caused by loss of
CC	proliferative capacity and overexpression of hypertensive and thrombotic
CC	factors in endothelial cells. Expanded populations of normal or
CC	genetically engineered rejuvenated cells could be used for autologous or
CC	allogeneic cell and gene therapy. They can also be used for prolonging
CC	the lifespan of a culture of normal cells or tissue being used to secrete
CC	therapeutic or other commercially significant proteins and products.
XX	
SO	Sequence 4027 BP: 674 A; 1361 C; 1277 G; 715 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4027;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

Qy	23	CCGCGCTTGAGCA	CAATGAC	CCGCGCCTCTCGTTG	CCCCCGGCGTCTCTGCTGCTGCG	82
Db	41	CCCGCCACCC	CCGCGATCGCGCGCTCCCGCTGCCAG	CCGTCGCTCCCGTCTGCTGCTGCG	100	
Qy	83	CAGCGATAC	CGGAGAGTGGCGCGCTGCGCA	CTTTGTGCGCGCGCTTGGGGCCCCGAGGG	142	
Db	101	CAGCCACTAC	CGAGGTGCTGCGCTGGGCA	GCTTCGTGCGCGCGCTTGGGGCCCCGAGGG	160	
Qy	143	CAGCGCGGTTGTGCA	ACCCGGGGA	CCGGAAGATCTACCGCA	TTTGGTTGGCCCAATGCT	202
Db	161	CTGGCGGCTGTGTAG	CGCGGGGAC	CCGGGCGCTTTCCGGCGCTGCTGGCCAGTGGCT	220	
Qy	203	AGTGTGCATGCA	TGGGGCTTCACAG	CTCCACCTGTCCGAC	CTTTCCTCCACCAAGTGC	262
Db	221	GGTGTGGCTGGCTTGGAC	GCACAGGCGCGCCCCCGCGCGCCCTCTCTTCGACAGGTGC	280		

263 ATCCCTGAAAGAGCTGTGGCCAGGGTTGTGAGAGACTCTCGGAGCGCAAGCAGAGAAA 322  
281 CTGCTGAGAGAGCTGTGTGGCCAGTGTCTGAGAGCTGTGCGAGCGGGCGCAAGAA 340  
323 CGTCTGGCTTTTGGCTTTGAGTGTCTTAAAGAGCCAGAGCGGGCTCTCCCATGGCCCTT 382  
341 CGTCTGGCTTTGGCTTTGAGTGTCTTAAAGAGCGGGCGGGCGGGCGGGCGGGCTT 400  
383 CACTAGTAGCGTGTGAGTACTTGTCCCAACTGTATTGAGACCTGCGTGTGAGTGG 442  
401 CACACAGCGTGTGAGTACTTGTCCCAACTGTATTGAGACCTGCGTGTGAGTGG 460  
443 TGCATGATGCTACTTGTGAGCGGAGTGTGGCGAGCACTGTCTGTCTGAGTGG 502  
461 GGGTGGGGCTGTGTGTGGCGCGGTGTGGCGAGCAGTGTGTGTGTGTGTGTGTGTGT 520  
503 CTGTGTCTTTATCTTCTGTGTGTCGCCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 562  
521 CTGTGTGTCTTGT 580  
563 GTACCAAAATTTGTGACCAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 622  
581 GTACCAAGT 640  
623 CGACCGGTGGCAGGAATTTCACTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 682  
641 GCGTGTGGGATGC-----GAAACGGGCTGTGAAACCATAG 673  
683 TCGCCAGGAAGACCGAAACCCCTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 742  
674 CGTCAGGGAGCGGGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 733  
743 TCTCACAGTACAAAGT 802  
734 CAGTGTGAGCGGAGT 793  
803 GGAGGGAGGACCCACAGCAGT 862  
794 GGAGCGGACCGCGT 853  
863 TCTGT 913  
854 TGACCGT 913  
914 AGGAAGGT 970  
914 GGT 973  
971 CTCACAT---CTCTGT 1027  
974 CCCCCATCCACATCGCGGCCACAGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033  
1028 TGAGACGACATTTCTTTTCTCAGGGGAGATGGCCAGAGCTGTAAACCCCTCATTT 1087  
1034 CGAGACCAAGCACTTCTTCTCTCTCAGGCA---CAAGGAGCAGTGTGTGTGTGTGTGT 1090  
1088 CTTACTCAGCAACTCTCAGCTTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147  
1091 CTTACTCAGCTCTCTGAGGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1150  
1148 TCTGT 1207  
1151 TCTGT 1210  
1208 ATACTGACAGT 1267  
1211 CTACTGGCAATTCGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1270  
1268 TGTGAGACTCTCAGGT 1327  
1271 CGGGGTGTCTCTCAAGACGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1330

1328 CT-----TGAACACAGGCC 1342  
1331 CTGTGCCCGGAGAAAGCCCCAGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1390  
1343 ACCGACCTTCATGGATTTTGTCTCCGCTGTGCACAGAGTCCCTTGGCAGGTATATGTTTCT 1402  
1391 CGGTCCGCTGTGTGAGTGTCTCCGCGACACAGCAGCCCTTGGCAGGTGTACGCTTCGT 1450  
1403 TGGGGCTGTCTCTGAAAGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1462  
1451 GGGGGCTGTCTGT 1510  
1463 CCGCTCTTTTAAAGAACTTAAAGAACTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1522  
1511 CCGCTTCTCAGNAACCAAGAGTTCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1570  
1523 GAGGAACTGT 1582  
1571 GAGGAGT 1630  
1583 GAAGGACGT 1642  
1631 GGT 1690  
1643 GTTGT 1702  
1691 GCACTGGT 1750  
1703 GAGCACTTCCAGAAAGACAGCTCTTCTTACCGTAAGAGTGTGTGTGTGTGTGTGTGTGT 1762  
1751 GACCACTTTCAAAGAAACAGCTCTTCTTACCGTAAGAGTGTGTGTGTGTGTGTGTGTGT 1810  
1763 GAGCACTTCCAGAAAGACAGCTCTTCTTACCGTAAGAGTGTGTGTGTGTGTGTGTGTGT 1822  
1811 AAGCACTGGAATCAGACAGCACTTGAAGAGGTGCAGTGTGTGTGTGTGTGTGTGTGTGT 1870  
1823 GGT 1882  
1871 GGT 1930  
1883 CAAGCCAAACCGCTGT 1942  
1931 CAAGCTGT 1990  
1943 GGGCAGAGGAGCAGGGCCAGCACTTCCAGCGTCTCAAGACTCTCTTTCAGCATGCT 2002  
1991 CGGCAGAGAAAGAGGGCCAGCGTCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2050  
2003 CAACCTATGAGCGGCAAAACATCTCTACCTTATGGGTCTTCTGTGTGTGTGTGTGTGTGT 2062  
2051 CAACCTACGAGCGGGCGCGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2110  
2063 CATCTACAGGACTGT 2122  
2111 TATCCAGAGGCTGT 2170  
2123 GATGTACTTTTGTAAAGCAGATGTGACCGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2182  
2171 GCTGTACTTTTGTCAAGT 2230  
2183 GGTGTAGT 2242  
2231 CACGGAGGT 2287  
2243 TGCAGT 2302  
2288 TCCGT 2347  
2303 CACCTCTCTGACCTTCCAGCCATATATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2362  
2348 TACCTTGT 2401  
2363 TGCCAGT 2422



2402 GACCAGCCGCTGAGGGATGCGGTGCTCATCGAGCAGAGCTCCTCCCTGAAATGAGGCAG 2461

2423 CAGCAGCCGTGTTGACTTCTCCCTGCACTTCCCTCGCTCAGCTGTCGTAAGAATGTGTGA 2482

2462 CAGTGGCCTCTTCGACGCTCTCTCACTGCTTCATGTGTCACACGCGGTGCGCATCAGGGG 2521

2483 CAGGTGCTATACGACATGTCGACGGGCATCCCCAGGGGCTCCAGCCTATCCACCCTGCTCTG 2542

2522 CAACTCTACGTCAGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2581

2543 CAGTCTGTGTTTCGAGACATGGAGAACAGCTGTTTGTCTGAGGTGACAGCGGATGGGTT 2602

2582 CAGCTGTGCTACGGCGACATGGAGAACAGCTGTTTGGCGGATTCGCGCGAGCGGCT 2641

2603 GCTTTTACGTTTGTGTGAATCTTCTGTGTTGTGACGCCCTCATCTTGGACCAAGCAAAAC 2662

2642 GCTCCTCGCTTGTGTGGATGATTTCTTGTGTGTGACACCTCATCCTCACCACGCGCAAAAC 2701

2663 CTTCTCTCAGCACCCCTGGTCCATGCGGTTCTGTAGTATGGGTGCATGATAAATCTGCAGAA 2722

2702 CTTCTCTCAGAACCTTGGTCCAGAGTGTCCCTGATATGCTTGGTGGTGAATCTGCGAA 2761

2723 GACAGTGTGAATCTTCCCTGTGTGAGACCTGTGATCCCTGGGTGTGAGCTCCATACAGCT 2782

2762 GACAGTGTGAATCTTCCCTGTAGAGACGAGGGCCTGGGTGGACGGCTTTTGTTCAGAT 2821

2783 GCGTGTCTACTGCTGTTTCCCTGGTGTGCTTGTCTGTGTCGTCGACACTCAGACTTTGGAGGT 2842

2822 GCCGGCCACGGCCTATTCCCTCTGTGGCGCTTGTCTGTGATACCCGACCCCTGGAGGT 2881

2843 GTTCTGTGACTACTCAGGTATTGCCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAG 2902

2882 GCAGAGCGACTACTCCAGCTATGCCCGACCTCCATCAGAGCCAGTCTCACCTCAACCG 2941

2903 TGTCTTCAAGCTGGGAGACCATGCGGACAGACCTCTGTGGTCTTGGGTGAGTG 2962

2942 CGCTCTTCAAGGCTGGGAGAACATGCGTCGCAAACTTTTGGGGTCTTGGCGCTGAAGTG 3001

2963 TCACGGTCTATTCTTAGACTTTCGAGGTGAAACAGCCTCCAGACGTCTGCATCAATATATA 3022

3002 TCACAGCCTGTTCTTGGATTTGCAGTGAACAGCCTCCAGAGGTGTGCACCAACATCTA 3061

3023 CAAGATCTTCTGTCTCAGCCCTACAGGTTCCATGCAATGTGATTTACGTTTCCTTTGA 3082

3062 CAAGATCTCTCTGCTGCAGGGCGTACAGGTTTCACGCAATGTGCTGCAGCTCCCATTTCA 3121

3083 CCAGCGTGTTAGAGAACTTCAATCTTCTGGGCATCATCTCCAGGCCAAGCATCCTG 3142

3122 TCAGCAAGTTTGGAGAACCCCAATTTTCTGCGGTCTATCTCTGACACGGCCTCCCT 3181

3143 CTGCTATGCTATCTCTGAAGGTCAAGATTCAGGAATGACACTAAAG-----GCCTC 3193

3182 CTGCTACTCCATCTCTGAAAGCCAGACGCAAGGATGTCTCGGGGGCCAGGGGCGCGCG 3241

3194 TGGCTCTTTTCCCTGTGAAGCGCACATTGGCTCTGTCTACAGGCTTCTCTCTCAAGCT 3253

3242 CGGCCCTCTGCCCTCCGAGGCCGTGCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3301

3254 GCGTGTCTATTTGTTCATCTAATAATGTCTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313

3302 GACTTCGACACCGTGTCACTAGTGCCTCTCTGGGGTCTCTCAGGACAGCCGACGCCA 3361

3314 GCTGTGCGGAAGCTCCACAGAGCGCAATGACCATCTTTAAAGCTGCAGCTGACCCAGC 3373

3362 GCTGAGTCCGAAGCTCCCGGGAAGACGCTGATGTCCTTGAGGCGCGAGCCACCCGCG 3421

3374 CCTAAGCACAGACTTTCAACACATTTGGACTAA 3407

3422 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3455

AA29388	AAA29388 standard; CDNA; 4027 BP.
ID	XX
XX	AAA29388;
XX	12-SEP-2000 (first entry)
DT	XX
DE	hEST2, a human telomerase catalytic subunit homologue cDNA.
XX	hEST2; telomerase; catalytic subunit; reverse transcriptase; life-span;
XX	retinoblastoma; p53; tumour suppressor; inhibitor; arteriosclerosis;
KW	proliferation; immortal; tumour therapy; macular degeneration; activator;
KW	INK4; ss.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key Location/Qualifiers
FT	CDS 57..3455
FT	/*tag= a
FT	/product= human_telomerase_catalytic_subunit
XX	XX
DN	W0200031238-A2.
XX	XX
PD	02-JUN-2000.
XX	XX
PF	24-NOV-1999; 99NC-US27907.
XX	XX
PR	25-NOV-1998; 98US-0109891.
PR	17-FEB-1999; 99US-0120549.
XX	XX
PA	(GENE-) GENETICA INC.
XX	XX
PI	Hannon GJ, Beach DH;
XX	XX
DR	WPI; 2000-400055/34.
DR	P-FSDB; AAY36566.
XX	XX
PT	New method for increasing the proliferative capacity of cell lines
PT	comprises administering agents reversibly activating telomerase
PT	activity and reversibly inactivating RB/INK4 and/or p53 pathways useful
PT	in treating age related diseases
XX	XX
PS	Claim 14; Page 112-116; 123pp; English.
XX	XX
CC	This cDNA, designated hEST2, is a human telomerase catalytic subunit
CC	homologue of yeast E8T2p and Eplotes p123. hEST2 is a member of the
CC	reverse transcriptase family of enzymes.
CC	The invention concerns methods and reagents for extending the life-span,
CC	e.g. the number of mitotic divisions, of a cell. The method relies on
CC	activation of a telomerase activity and inhibition of one or both of a
CC	retinoblastoma (RB)/INK4 pathway or a p53 pathway. Phosphorylation of RB
CC	by cyclin-dependent kinases, cdk4 and cdk6, releases the cells into the
CC	division cycle. Binding of INK4 family members, e.g. the tumour
CC	suppressor p16INK4a, inhibits kinase activity and results in growth
CC	arrest. Rb inactivators can selectively and reversibly inactivate an
CC	RB/INK4 pathway, especially an RB/p16INK4a pathway. The oncoprotein MDM2
CC	is a cellular inhibitor of Rb/E2F function and the p53 tumour suppressor
CC	can also be used in the methods. Other molecules which can be used
CC	include cdk4 or cdk6 mutants. In particular, a cdk4 mutant is one which
CC	differs from at one or more of residues K22, R24, H95 and/or D97.
CC	Additional constructs include a papilloma virus E7 protein, or other
CC	viral oncoprotein which bypasses Rb and/or p53. Antisense constructs of
CC	the Rb and p16INK4a genes may also be used. The methods are useful for
CC	increasing the proliferative capacity of cells. The cells are
CC	subsequently of use in pharmaceutical and cosmetic preparations used to
CC	treat conditions related to (premature) ageing, e.g. macular degeneration
CC	and arteriosclerosis. The cells can also be used to replace tumour cells
CC	lines in vitro and for studies on biochemical and physiological aspects
CC	of growth and differentiation. Long lived (immortal) cells could also be
CC	of use in the production of normal or genetically engineered
CC	biotechnology products.
XX	XX
SQ	Sequence 4027 BP; 674 A; 1361 C; 1277 G; 715 T; 0 other;

ery Match 45.8%; Score 1500.8; DB 21; Length 4027;  
 et Local Similarity 68.9%; Pred. No. 0;  
 tches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGAACCAATAGACCCCGCTCTCGTGGTCCCGCGGTGCGTCTCTGTGGG 82  
 41 CCGGCGCAACCCCGCGATGCGCGCTCCCGCTCGGAGCGTGGCTCCCTGTGG 100  
 83 CAGCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGGGCGCTGGGGCCGAGGG 142  
 101 CAGCGCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGGGCGCTGGGGCCGAGGG 160  
 143 CAGGCGGCTTGTGCAACCCCGGAGCCCGGAGATCTACCGCACTTTGGTTGCCCAATGCCCT 202  
 161 CTGGCGGCTGTGCGAGCGCGGGAGCCCGCGGCTTTCCGCGGCTGTGGCGCTGTGCCT 220  
 203 AGTGTGCTATGACTGGGCTCAGAGCTCCACCTGCGGACCTTTCTTCCCTCAGCAGGTGC 262  
 221 GGTGTGCGTGGCTGGGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 280  
 263 ATCCCTGAAAGCTGTGGCGAGGTGTGCGAGACTCTGCGAGCGCAACGAGAGAA 322  
 281 CTGGCTTGAAGAGCTGTGGCGCGGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGAGAA 340  
 323 CGTGTGCTTTTGGCTTTGAGCTGTAAACGAGCGGAGGCGGCGCTCCCATGSCCTT 382  
 341 CGTGTGCGCTTGGCTTTGGCTTTGAGCTGTGAGCGGCGCGCGCGCGCGCGCGCGCGCT 400  
 383 CACTAGTAGCGGTGCTAGCTTGTGCGCAACCTTTATGAGACCTGCTCTAGTGG 442  
 401 CACCAACGCGTGGCGAGCTACTCTGCGCAACACCGGTGACCGACCACTGCGGGGAGCGG 460  
 443 TGCATGATGCTACTTGTGAGCGGAGTGGGAGCGACCTGTGCTTACCTGTGCGCACA 502  
 461 GCGTGGGCGCTGTGTTGCGCGCGTGGCGAGCGAGCTGTGCTTCACTGTGCGCAG 520  
 503 CTGTGCTTTTATCTTCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 562  
 521 CTGCGCGCTTCTTGTGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 580  
 563 GTACCAAAATTTGTGCGCAGCGATATCTGCGCGCTCTGTGCTGCTAGTTTACAGGCGCAC 622  
 581 GTACAGCTGCGCGCTGCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 640  
 623 CCGACCGGTGGCGAGAAATTTCACTAACCTTAGGTGTTTCAACAGATCAAGAGAGTAG 682  
 641 GCGTCTGGGATGC-----GAAACGCGCGCTGGAAACCATAG 673  
 683 TCGCAGGAGACCGAAACCCCTGGCGCTTGCCTATCTCGAGGTACAAGAGGCTCTGAG 742  
 674 CGTAGGAGGCGCGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 733  
 743 TCTCAGGATACAGGTGCGCTTACGTAAAGGCGGAGATGCTATCTGTGCGGAGAGT 802  
 734 CAGTGCAGCGCAAGTCTGCGCTTGCCTCAAGAGCGCGAGGCGTGGCGCTGCGCGCTGAGCC 793  
 803 GGAGGAGGACCCACAGCGAGGTGCTACCAACCCCATCAGGCAATCATGGGTGCCAG 862  
 794 GGAGCGGAGCGCGCTTGGGAGGCGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 853  
 863 TCGTGTGCGCTCCCGGAGGTGCTTACT-----GCAGAGAAAGATTGTCTTTCTAA 913  
 854 TGACCGTGTGTTCTGTGTGTGCTACCTGCGCAGACCGCGCGAGAACCCCTCTTTTGA 913  
 914 AGGAAGGTGCTTGAACCTGAGTCTCTC---TGGGTGCGTGTGCTGTAAACACAGGCCAG 970  
 914 GGTGTGCGCTCTTGTGACGCGCGCACTCCACCCCATCCGTCGCGCGCGCGCGCGCGCG 973  
 971 CTCCACAT---CTCTGCTGTACCAACCGCGCAAAATGCTTTTTCAGTCTCAGGCGCATTTAT 1027  
 974 CCCCCATCCATCGGCGCACACAGCTCTCTGGGACACGCGCTTGTCCCCCGGTGTACGC 1033

QY 1028 TGAGACAGACATTTCTTTTACTCCAGGAGATGCGCAAGAGCGTCTAAACCCCTCATTT 1087  
 DB 1034 CGAGACCAACACATTTCTTCTACTCTCTCAGCGA---CAAGGAGAGCTGCGCGCTCTT 1090  
 QY 1088 CTTACTCAGCAACTCCAGCTTAACCTTGACTGCGGCGCAGAGACTGGTGGAGATCATCTT 1147  
 DB 1091 CTTACTCAGCTCTCTAGGCGCAGCTGACTGCGCTCGGAGCGCTGCTGGAGACCATCTT 1150  
 QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGAACCATCTCTGAGGACACACCGCTCTATCGGCTG 1207  
 DB 1151 TCTGGGTTCCAGGCTTGGATGCGAGGACTCTCCCGAGTTGCCCCGCTGCCCGAGCG 1210  
 QY 1208 ATACTGAGATGCGGCGCTCTGTTTCCAAAGCTGTGTTGTAACCATGAGAGTGCATA 1267  
 DB 1211 CTACTGCAAAATGCGGCGCTGTTTCTGAGCTGCTTGGAAACCAAGCGCAGTGGCCCTA 1270  
 QY 1268 TGTGAGACTCTCTCAGGTCAATGAGGTTTGGACAGCAACCAACAGGTGACAGATGC 1327  
 DB 1271 CGGGTGTCTCTCAAGACGACTGCGCGCTGAGCTGCGGTACCCAGCAGCGCGGT 1330  
 QY 1328 CT-----TGAAACACCAAGCGCC 1342  
 DB 1331 CTGTGCGCGGAGAGCGCCAGGCTCTGTGGCGCGCCCGGAGGAGGACACAGACCC 1390  
 QY 1343 ACCGCACTCATGATTTGCTCGCTGCGACAGCTGCTGCGAGGTATATGTTTCT 1402  
 DB 1391 CCGTGGCTGTGCTGCTGCGCGCAGCAGCAGCGCTGCGAGGTGTACGGCTTCT 1450  
 QY 1403 TCGGCGCTGCTCTGCAAGGTGTGCTGCTGCTGCGGTACAGGCAACAAGAGCG 1462  
 DB 1451 CCGGCGCTGCTGCGCGGCTGTGCGCGCGCTCTGGGCTCCAGGCAACAAGAG 1510  
 QY 1463 CCGCTTCTTAAAGACTTAAAGAGTTCATCTGTTGGGAAATACGGCAAGCTATCAT 1522  
 DB 1511 CCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGAAAGCATGCCAAGCTCTCGT 1570  
 QY 1523 GCAGAACTCATGTGGAAGATGAAGTAGAGGATTCGCTGCGCGAGCGCGCGCG 1582  
 DB 1571 GCAGAGCTGAGTGGAGATGAGCGTGGCGGCTGCGTGGCTGCGGAGCGCGCG 1630  
 QY 1583 GAAGAGCGCTGCTCCCGCTGCGAGCAAGCTCTGAGGAGAGGATCCTGGTACGTTCT 1642  
 DB 1631 GGTGTGCTGTTTCCGCGCGCAGAGACCGCTCTGCGTAGGAGATCCTGGCGAAGTCTCT 1690  
 QY 1643 GTTCTGCTGATGACACATACGTTGTAAGTGTAGTCTTAGTCTATCTTTTATCATCACA 1702  
 DB 1691 GCATGCTGATGAGTGTGACCTGCTGAGCTGCTGAGTCTTTTATGTCACGA 1750  
 QY 1703 GAGCACATTCAGAGAAACAGGCTCTTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
 DB 1751 GACCACTTCAAAAGAAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1810  
 QY 1763 GAGATTTGAGTCAAGGCAACCTTGAGAGTGGGCTACGGAGCTGTCAAGAGAGA 1822  
 DB 1811 AAGCATTTGAATCAGACAGCACTTGAAGAGGTTGAGCTGCGGAGCTGTGGAAGCAGA 1870  
 QY 1823 GGTCAAGCATCACAGACACCTTGGCTAGCATGCGCATCTGCGAGCTGCGTTATCCC 1882  
 DB 1871 GGTCAAGGAGATCGGAGAGCAGGCGCGCTGCTGAGCTGCGAGCTCGGTTATCCC 1930  
 QY 1883 CAAGCCCAACGCGCTGCGGCCCATTTGTGAACATGATGATFATAGCATGGTACAGAGCTTT 1942  
 DB 1931 CAAGCTGACGCGGCTGCGGCGGATTTGTGAACATGAGTACGTCTGGGAGCGCAAGACGTT 1990  
 QY 1943 GGGCAGAGGAGAGCGAGCGCGCATTTCAACAGCTCTCAGAGCTCTCTTCCAGCATGCT 2002  
 DB 1991 CCGCAGAGAGAAAGAGGCGCGAGGCTCTCACCTGAGGTTGAAGGCACTGTTTACGCTGCT 2050  
 QY 2003 CAATATGAGCGGCAAAACATCTTATGAGGCTCTTCTGTACTGGGTATGAATGA 2062  
 DB 2051 CAATACAGCGCGCGCGCGCGCTCTTCTGGGCGCTCTGTGCTGGGCTGGACGA 2110  
 QY 2063 CATCTACAGGACCTGGCGGCGCTTTGTGCTGCGTGTGCTGCTGCGACACACCCAG 2122

```
2111 TATCCAGAGGCTGGCGACCTTCGTGCTGGGCGCCAGGACCCGCGCCCTGA 2170
2123 GATGTATTTTAAAGCAGATGTGACCGGGGCTATGATGCCATCCGCCAGGTAAGCT 2182
2171 GCTGTATTTGTCAAGTGGATGTGAGGGGCGGTACGACACCATCCGCCAGGACAGCT 2230
2183 GGTGGAGGTGTGGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242
2231 CACGGAGGTATCGCCAGCATATCA---AACCCAGAACACGTACTCGGTGCGTGGTA 2287
2243 TGCAGTGGTCCGAGAGATAGCCAAAGCCCAAGTCCCAAGTCCCTTTAGGACACAGTCA 2302
2288 TGGGTGGTCCAGAGGCGCCATGGGCACTCGGCCAAGGCTTCAAGAGCCACGCTC 2347
2303 CACCTCTCTGACTCCAGCATACATGGGCGAGTTCCTTAAGCATCTGCAGGATCAGA 2362
2348 TACCTTGACAGACCTCCAGCGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2401
2363 TGCAGTGCACCTCAGGAACCTCGGTGTCTATCAGCAGAGCATCTCTATGATCAGAGCAG 2422
2402 GACAGCCGCTCAGGGATGCGGTCTATCAGCAGAGCTCTCTCCCTGATGAGCCAG 2461
2423 CAGCAGCTGTTTGAATCTTCTGCACTTCTGCGTCAAGTGTGCGTAAAGATTGGTA 2482
2462 CAGTGGCTCTTCGAGCTCTTCTAGCTTCTATGTGCCACACGCGCTGCGCATCAGGG 2521
2483 CAGTGTATATACAGTGCAGGCGATCCCGAGGCTCCAGGCTTATCCCTGCTCTG 2542
2522 CAGTCTAGTCCAGTGCAGGGATCCCGCAGGCTCATCTCTCAGCTGCTCTG 2581
2543 CAGTCTGTGTTTCGAGACATGAGAACCAAGTGTGTTGAGGTGCGACGCGGATGGTT 2602
2582 CAGCTGTGCTACGGCGACATGAGAACCAAGTGTGTTGCGGGATTCGCGGAGCGCT 2641
2603 GCTTTACGTTTGTGATGACTTCTGTGTGTGCGCTCACTTGGACCAAGCAAAAC 2662
2642 GCTCTCGTGTGGTGAATTTCTGTGTGTGACACCTCACTTCCCTCAGCGGAAAC 2701
2663 CTTCTCAGCACCCTGCTCATCGCTTCTGATGATGGTGCATGATAAATTCAGAA 2722
2702 CTTCTCAGCACCCTGCTCGAGGTGCTGATGATGCTGCTGCTGATGCTGCGAA 2761
2723 GACAGTGTGAATCTCCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782
2762 GACAGTGTGAATCTCCTGTGAGACGAGGCGCTTGGTGGACGCTTGTTCAGAT 2821
2783 GCCTGCTCACTGCTGTTTCCCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842
2822 GCGGCGCCAGGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2881
2843 GTTCTGTACTCAGGTTATCCCGAGCTCAATTAAGACGAGCCTCACCTCCAGAG 2902
2882 GCAGAGGACTACTCCAGCTATCCCGGACCTCCATCAGAGCAGTCTCACCTTCAAC 2941
2903 TGTCTTCAAGTGGGAGACCATGCGAACAGCTTCTGCTGCTGCTGCTGCTGCTGCT 2962
2942 CGGTTTCAAGCTGGGAGAACATGCGTCCAAACCTTTGGGCTTTCGCGCTGAAGT 3001
2963 TCAGGCTTATTTTACACTTGAAGTGAACAGCTTCCAGAGCTGCTGATCAATATATA 3022
3002 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCTTCCAGAGCTTCCAGACCATCTA 3061
3023 CAAGATCTTCTGCTTCCAGGCTACAGTTCATGATGTGATTCAGTTCCTTTGA 3082
3062 CAAGATCTTCTGCTGAGGCTACAGTTCATGATGTGCTGCTGCTGCTGCTGCTGCT 3121
3083 CCAGGCTTTAGGAGAACCTCAATTTCTTGGGCTATCTTCCAGCAGCAAGTCTCTG 3142
3122 TCAGCAAGTTTGGAGAACCCCAATTTTCTGCGGCTCATCTGACACGCGCTCCT 3181
3143 CTGCTATGCTATCTGAGTGCAGATCCAGGATCAGGATGACATTAAG-----GCCTC 3193
```

```
Db 3182 CTGCTACTCCTCTGAAAGCCAGGAGTGTGCTGGGGCGCAGGCGCCGC 3241
QY 3194 TGGCTCTTTCTCTGAGCGCCACATTGGCTCTGTACAGGCTTCTGCTCAAGCT 3253
Db 3242 CGGCGCTCTGCGCTCCGAGGCGGTGCACTGGCTGTGCCACCAAGCATTCCTGCTCAAGCT 3301
QY 3254 GGTCTCTCATTTCTGTCATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313
Db 3302 GACTCGACACCGTGTACCTAGTGCCTCTCTGGGTCACTCAGGACAGCCGAGCGCA 3361
QY 3314 GCTGTGCGGAGTCCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGACCCAGC 3373
Db 3362 GCTGAGTCGGAAGTCCCGGGAGAGGCTGACTGCTGAGGCGGAGCCAGCCAGCCGCGC 3421
QY 3374 CCTAAGCAGACACTTTCAGACCATTTTGGACTTA 3407
Db 3422 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3455
```

## RESULT 14

```
AAV72117
ID AAV72117 standard; cDNA; 4042 BP.
XX AC AAV72117;
XX DT 24-MAY-1999 (first entry)
XX XX Human catalytic telomerase sub-unit cDNA.
XX DE Human; catalytic telomerase subunit; therapy; diagnosis; hTC; assay;
XX KW modulator; treatment; inhibit; cellular disorder; death; defect; cancer;
XX KW ageing; antisense; neoplastic cell; telomerase-related condition;
XX XX tumour cell; ss.
XX OS Homo sapiens.
XX FH Key
XX CDS Location/Qualifiers
FT FT 63..3461
FT FT /*tag= a
FT FT /product= "catalytic telomerase subunit"
```

WO9859040-A2.

30-DEC-1998.

09-JUN-1998; 98WO-EP03468.

14-APR-1998; 98DE-1016496.

20-JUN-1997; 97DE-1026329.

26-MAR-1998; 98DE-1013274.

(FARB ) BAYER AG.

Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;

WPI; 1999-081276/07.

P-PSDB; AAW90251.

New catalytically active subunit of human telomerase - used in the modulation of telomerase activity, particularly for treating cancer and ageing

Claim 4; Fig 1; 76pp; German.

This sequence encodes a novel human catalytic telomerase sub-unit (hTC). The encoded protein can be used in screening assays to identify modulators of telomerase and to treat or inhibit cellular disorders, death, defects and/or other pathological processes involving telomerase, particularly cancer and ageing (also suitable for this are agents that stimulate, inhibit or mimic the activity of the subunit). Antisense nucleic acids inhibit telomerase action (by binding to specific mRNA), particularly in neoplastic cells and may be expressed in vivo. Antibodies and fragments of the protein, used as probes or primers, are used to

diagnose telomerase-related conditions (especially neoplasia) by (i) detecting abnormal levels of the subunit protein in body fluids or tissues or (ii) by measuring the amount of the encoding nucleic acid. Expression of the nucleic acid encoding the subunit mRNA is confined to tumour cells, in contrast to the ubiquitous expression of the telomerase RNA subunit.

Sequence 4042 BP; 584 A; 1364 C; 1277 G; 717 T; 0 other;

very Match 45.8%; Score 1600.8; DB 20; Length 4042;  
est Local Similarity 68.9%; Pred. No. 0;  
atches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

```
23 CCGGCGCTTGAGACAAATGACCGCGCTCTCGTCCCGCGGTGCGTCTCTCTCGG 82
47 CCGGCGCACCCCGCGATGCGGGCGTCCCGCTGCGGACCGTGGCTCTCTCTGG 106
83 CAGCCGATACCGGAGGTGTGGCGGTGGCAACTTTGTGCGGCGCTGGGGCCGAGG 142
107 CAGCCTACTCCGCGAGGTGTGGCGGTGGCAAGTTCTGTGCGGCGCTGGGGCCGAGG 166
143 CAGCGGCTTTGTCAACCCCGGACCCGAGATCTACCGCACTTTGGTGGCCCAATGCT 202
167 CTGGCGCTCTGTGAGCGCGGACCCCGCGGTTCCTCCGCGCTGTGGCCCAATGCT 226
203 AGTGTGATCATCTGGGGCTCACAGCTCTCCACTCGCGCACTTTCTTCCACAGGTCT 262
227 GGTGTGCTGCTCGCTGGACGACGCGCGCGCCCGCGCGCTCTCTCCGCGAGGTCT 286
263 ATCCCTAAAGAGCTGTGGCGAGGTGTGGAGAGACTCTGGAGCGCAACGAGAA 322
287 CTGCTTGAAGAGCTGTGTGGCGCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCAAGAA 346
323 CGTCTCGCTTTTGGCTTTGAGCTGTCTTAACAGGCGCAGAGCGCGCTCTCCATGCGCTT 382
347 CGTCTCGCTTTTGGCTTTGAGCTGTCTTAACAGGCGCAGAGCGCGCTCTCCATGCGCTT 406
383 CACTAGTAGCGTGTAGTACTTGTGCGCGCAACACTGTATTATGAGACCTCGTGTGAGTGG 442
407 CACCACAGCGTGTGCGAGTACTTGTGCGCGCAACACTGTATTATGAGACCTCGTGTGAGTGG 466
443 TGCATGATGCTACTGTGTAGCGCGAGTGGCGCGAGCTGTGTGCTACTCTCTGGACA 502
467 GCGTGTGGGCTGTGTGTGCGCGTGTGGCGAGAGTGTGTGCTACTCTCTGGACG 526
503 CTGTGCTTTTATCTTGTGTGCGCGCGAGTGTGTGCTTACAGGTGTGTGTGCTTCTCCCT 562
527 CTGCGCGCTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 586
563 GTACCAATTTGTGCCACCGATATCTGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 622
587 GTACCAAGCTGTGCGCTGTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 646
623 CCGACCGCTGGGCGAGAAATTCATACTAGTGTCTTACCAAGATCAAGAGAGTAG 682
647 GCGTCTGGATGCG-----GAAAGCGCTGTGAACCATAG 679
683 TCGCGAGAGACCGAAACCCCTGGCGCTTGGCATCTCGAGGTACAAAGAGCATCTGAG 742
680 CGT CAGGAGGCGCGGCTTCCCTTGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 739
743 TCTCACCAGTACAGTGTGCTTCTCAGTAAAGAGCGCAGATGCTATCTGTCTCCGAGAGT 802
740 CAGTGTCCAGCGGAGTCTGCGTTTCCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 799
803 GGAGAGGAGACCCACAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 862
800 GGAGCGAGCGCGCTTGGGCGAGGCTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 859
863 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 913
860 TGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 919
```

```
QY 914 AGAAGAGTGTCTGACCTGAGTCTCTC---TGGTGGTGTGTGTGTAAACACAGCCAG 970
DB 920 GGTGGGCTCTCTGGACGCGCACTCCACCCATCCGTGGCGCGCGACACACAGCGGG 979
QY 971 CTCACAT---CTCTGCTGTACACCCCGCAAAATGCCCTTTTTCAGTCTAGGCCATTTAT 1027
DB 980 CCCCCATCCATCGCGCGCACACAGTCCCTGGGACACGCTTGTCCCCCGGTGTACGC 1039
QY 1028 TGNAGCAGACATTTCTTACTCCAGGGGAGTGGCCAAAGAGCTCTAAACCCCTCAT 1087
DB 1040 CGAGACCAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1096
QY 1088 CTACTACAGCAACCTCCAGCTTAATTTGACTGGGCGCAGGAGACTGTGTGAGATCATCT 1147
DB 1097 CTTACTCAGCTCTCTGAGGCCCGCAGCTGACTGGGCTCGGAGGCTCGTGGAGACATCT 1156
QY 1148 TGTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCGTCTATGCGGTG 1207
DB 1157 TCTGGGTTCAGGCGCTTGGATGCCAGGAGCTCCCGCAGGTTCGCCCGCTGCCCCCAGCG 1216
QY 1208 ATACTGGCAGATGCGGCGCTCTTCCAAACAGCTGTGTGTGAACCATGCAGAGTGGCAATA 1267
DB 1217 CTACTGGCAATGCGGCGCTCTTCTGAGCTGTCTTGGGAACCAACGCGCAGTGGCCCTA 1276
QY 1268 TGTGACATCTCAGTCAATTCAGTGTTCGAAACAGCAAAACCAACAGGTGACAGATGC 1327
DB 1277 CGGGTGTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1336
QY 1328 CT-----TGAAACACAGCGCC 1342
DB 1337 CTGTGCGCGGAGAGCGCCCGGCTCTGTGGCGCGCGCGAGGAGGAGACACAGACCC 1396
QY 1343 ACAGCACTCATATGATTTGCTCCGCTGCAACAGAGTCCCTGGCAGGTATATGTTTCT 1402
DB 1397 CCGTCCCTGTGTGAGCTGCTCCGCGACACACAGCAGCCCTTGGCAGGTGTACGCTTCT 1456
QY 1403 TGGGCGCTCTCTGCAAGTGTGTCTGTAGTCTCTGSGGTACCGGACAAATGAGCG 1462
DB 1457 GCGGCGCTTCCCTGCGCGGCTGTGTGCGCGAGGCTCTGSGGTCCAGGCAACAGAACG 1516
QY 1463 CCGCTTTCTTTAAGAACTTAAAGAGTTCATCTGTTGGGAAATACCGCAAGCTATCAT 1522
DB 1517 CCGCTTCTCTAGGAAACACCAAGATTCATCTCCCTGGGAAAGCATGCAAGCTCTGCT 1576
QY 1523 GAGGAACTCATGTGTGAGATGAAAGTAGAGTTCACCTGGCTCCGCGCAGCGCCCGG 1582
DB 1577 GCAGGAGCTGAGCTGGAAGATGAGCTGCGGAGCTGCGTGTGGCTGCGCAGGCGCCAGG 1636
QY 1583 GAAGGACCGTGTCCCGCTGACAGACACCGTCTGAGGAGAGGATCTCTGGCTACGTTCT 1642
DB 1637 GGTGGCTGTGTCCGCGCGCAGACACCGTCTGCTGAGGAGATCTCTGGCCAAAGTCTCT 1696
QY 1643 GTTCTGGCTGATGAGACATACGTGTGATCAGCTGTAGGTTCATCTTTTACATCACA 1702
DB 1697 GCATGGCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1756
QY 1703 GAGCATTCTCAGAGAAACAGGCTCTTCTTACCGTAAAGAGTGTGTGGAGCAAGCTGCA 1762
DB 1757 GACCACTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTGTCA 1816
QY 1763 GAGCATTGAGTCAAGCAACACTTGTGAGAGTGTGGGTACGGGAGCTGTCAAGAGGA 1822
DB 1817 AAGCATTGGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGAGAGCTGTGGAAGCAGA 1876
QY 1823 GGTGAGCATCAGGACACCTGCTAGCATGCCATCTGCGAGCTGCGCTTCATCCC 1882
DB 1877 GGTGAGCAGCATCGGGAAGCAGCGCGCTCTGTGCTGAGCTTCCGCTTCATCCC 1936
QY 1883 CAAGCCAAACGCGCTCGGCGCCCATTTGAAATAGTATAGCATGGGTACAGAGCTTT 1942
DB 1937 CAGCTGAGCGGCTGCGCGCGATTTGTGAACATGAGCTACGTCGTGGGAGCGAGACGTT 1996
QY 1943 GGGCAGAGAGAGCAGCGCCCGCATTTACCCAGCGCTCTCAGAGCTCTCTTCAGCATGCT 2002
```

Db	3058	CAAGATCCTCTGCTGAGGCGTACAGGTTTACGCAATGTGTGCTGAGCTCCCATTTCA	3122
QY	3093	CCAGCGTGTAGGAAGAACTCACAATCTTTTCTGGGCATCATCTCCAGCCAAAGCATCTGT	3142
Db	3128	TCAGCAAGTTTGAAGAATCCCCACATTTTTCCTGGCGCTCATCTCTGACACGGCTCCCT	3187
QY	3143	CTGCTATGCTATCTCTGAAGTCAAGAAATCCAGGAATGACACTTAAG-----GCTC	3193
Db	3198	CTGTACTCCATCTCTGAAGCCAAAGAACCCAGGGATGTGGTGGGGCCAAAGGCGCGC	3247
QY	3194	TGGTCTCCTTTCTCTGAAGCCGACATATGGCTCTGCTACCAAGGCTTCTGTCTCAAGCT	3253
Db	3248	CGGCGCTCTGCGCTTCGAGGCGGTGCACTGGTGTGTCACCAAGCATTTCTGTCTCAAGCT	3307
QY	3254	GGCTGCTCATCTGTCTATCTACAAATGTCTCTCTGGGACCTCTGAGGACAGCCCAAACT	3313
Db	3308	GACTCGACACCGTGTCACTTACCTTACGTGCCACTCTCTGGGTCACTCAGGACGCCACGCA	3367
QY	3314	GCTGTGCGGAAGCTCCGAGGCGGACAATGACCAATCCTTAAAGCTGAGCTGACCCAGC	3373
Db	3358	GCTGAGTCGGAGCTCCCGGGACGACGCTGACTGCCCTGGAGGCGCGCAGCAACCCGGC	3427
QY	3374	CCTAAGCACAGACTTTCAGACCATTCTTGGACTAA	3407
Db	3428	ACTGCTCCTCAGTACTCAAGACCATCTCTGGACTGA	3461
RESULT 15			
ABL53711	ID ABL53711 standard; cDNA; 4070 BP.		
XX	XX		
AC	ABL53711;		
XX	XX		
DT	17-JUN-2002 (first entry)		
DE	Human telomerase catalytic subunit hTERT cDNA.		
XX	XX		
KW	hTERT; telomerase; reverse transcriptase; immortalisation; human;		
KW	vaccine; enzyme; gene; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FN	WC200216555-A2.		
XX	XX		
PD	28-FEB-2002.		
XX	XX		
PF	17-AUG-2001; 2001WO-GB03726.		
XX	XX		
PR	17-AUG-2000; 2000GB-0020246.		
PR	17-AUG-2000; 2000US-225734P.		
XX	XX		
PA	{UYWA-} UNIV WALES COLLEGE OF MEDICINE.		
XX	XX		
PI	Jones CJ, Kipling DG, Wilkinson G, McSharry B, Skinner JW;		
XX	XX		
DR	WPI; 2002-315462/35.		
XX	XX		
PT	Novel hTERT-immortalized cell line (human telomerase reverse		
PT	transcriptase) useful for human vaccine production and preparation of		
PT	antigen, such as a virus or virus-derived agent -		
XX	XX		
PS	Example 1; Fig 1; 64pp; English.		
XX	XX		
CC	The present sequence is that of hTERT cDNA in plasmid pGRN121.		
CC	hTERT is the catalytic subunit of human telomerase. Claimed		
CC	immortalised cell lines for use in vaccine production are adapted		
CC	to express hTERT. Suitable cell lines comprise human diploid		
CC	fibroblasts, e.g. WEC-5 or WI38 cells, transfected with hTERT cDNA		
CC	or infected by a retrovirus carrying hTERT cDNA, and are capable		
CC	of supporting antigen production. A method for preparing such cell		
CC	lines using recombinant techniques is provided. The cell lines		
CC	are also used as a diagnostic test for the presence of a virus,		
CC	such as human cytomegalovirus, and to determine the efficacy of		

antiviral agents by testing the capability of a modified virus containing a reporter gene to infect the cells. The cell lines have the ability, which can be impaired in cell lines immortalised by other methods, of being able to support viral replication. The cells remain morphologically suitable for viral/vaccine cultivation.

Sequence 4070 BP; 670 A; 1383 C; 1296 G; 721 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 24; Length 4070;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

```
23 CCGCGCTTGAGCAATGACCCGCGTCTCGTTGCCCCGCGTGGCTCTCTGCTGG 82
60 CCGGACACCCCGGATGCGCGGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGG 119
83 CAGCCGATACCGGAGGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGGCCCGAGGG 142
120 CAGCCACTACCGGAGGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGGCCCGAGGG 179
143 CAGCGGCTTGTGCAACCCGCGGACCGAGATCTACCGACTTTGTTGTTGCCCAATGCT 202
180 CTGGCGCTTGTGCAACCCGCGGACCGCGGCTTTCCGCGCTGTTGCGCGCTGCT 239
203 AGTGTGATGCACTGGGCTCACAGCTTCAACCTTCCAGCTGCGGCTTTCTTCCACAGGTG 262
240 GGTGTGCTGCTGGAGCGACGCGCGCCCGCGCGCTTCCCTTCCGCGAGGTGTC 299
263 ATCCCTGAAGAGCTGGCGGAGGTGTCAGAGACTTGGAGCGCAACAGAGAA 322
300 CTGCTGAAGAGCTGGTGGCGCGAGTCTCAGAGGCTGTGAGGCGCGCGAGAA 359
323 CTGCTGGCTTTGGCTTTGAGCTGCTTAAAGCGCCAGAGCGCGGCTTCCCTTGGCTT 382
360 CTGCTGGCTTTGGCTTTGAGCTGCTTAAAGCGCGCGGCTTCCCTTGGCTT 419
383 CACTAGTAGCTGCTAGTACTTGGCCCAACACGTATTTAGAACCTGCTGCTAGTGG 442
420 CACCAACGCTGGCGAGCTACTTGGCCCAACACGTATTTAGAACCTGCTGCTAGTGG 479
443 TGCAATGATGCTACTTGTAGCGAGTGGCGAGCACTGCTGCTGCTGCTGCTGCTGCA 502
480 GCGTGGGGCTGCTGCTGCGCGGCTGGCGAGCACTGCTGCTGCTGCTGCTGCTGCTG 539
503 CTGCTGCTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
540 CTGCGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
563 GTACCAATTTGTCACACGATATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
600 GTACCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
623 CCGACCGCTGGCAGGAATTTCACTAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
660 GCGTCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
683 TCGCAGGAGACCGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
693 CBTGAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
743 TCTCAGCAGTACAGTGTGCTTCACTAAGAGCGCAGATGCTATCTGCTGCTGCTGCTGCT 802
753 CAGTCCAGCCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 812
803 GAGAGGAGGAGCCCAAGCAGGCTGCTACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
813 GAGAGGAGGAGCCCAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872
863 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
873 TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 932
```

```
QY 914 AGGAAGGTCTGTGACTGTGACTCTCTC---TGCGTGGTGTGCTGTAAACAAGCCCGAG 970
Db 933 GGGTGGCTCTCTGGACGGCCACTTCCACCATCTGGTGGGCCCGCAGCACCACGCGGG 992
QY 971 CTCACAT---CTCTGCTGTACACCCCGCAAAATGCTTTTCAGCTCAGGCCATTAT 1027
Db 993 CCCCCATCCACATCGCGGCCACAGTCTCTGGACACGCTTGTCTCCCGGTGTACGC 1052
QY 1028 TGAGACCAGACATTTCTTTTACTTCCAGGGGAGATGGCCAGAGCGTCTAAACCCCTCAT 1087
Db 1053 CGAGACCAGACATTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1109
QY 1088 CTTACTCAGCAACTCTCAGCTTAATTTGATGCGGCGCAGGAGACTGGTGGAGATCATCT 1147
Db 1110 CTTACTCAGCTCTCTGAGGCCAGCTGACTGCGCTGCGAGGCTCTGCGAGACATCTT 1169
QY 1148 TCTGGGCTCAAGGCTTAGACATCAGACCACTCTTGCAGGACACACCGCTCTATCGCTG 1207
Db 1170 TCTGGGTTCCAGGCCCTTGGATGCCAGGACTCCCGCAGTTGCGCCGCTTCCCGCAGG 1229
QY 1208 ATACTGCAGATCGGCGCTCTTCCAAACAGCTGCTGTGTGAACCATGACAGGTGCCAATA 1267
Db 1230 CTACTGCAATGCGGCGCTCTTCTGAGGCTGCTTGGGAACCGCAGCTGCGCTT 1289
QY 1268 TGTGAGCTCTCTCAGTCACTTGCAGGTTTGCAGAGCAACCAACAGGTGACAGTGC 1327
Db 1290 CGGGTCTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGAGCGCTG 1349
QY 1328 CT-----TGAAACACAGCC 1342
Db 1350 CTGTGCGCGGAGAGCCCGCAGGCTCTGTGGGCGCCCGAGGAGGAGACACAGACCC 1409
QY 1343 ACCGACCTCATGATTTGCTCCGCTGACAGAGTCTCTGCGAGGATATATGTTTCT 1402
Db 1410 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469
QY 1403 TCGGCGCTGCTCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
Db 1470 GCGGCTGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 1529
QY 1463 CCGTCTCTTAAAGAACTTAAAGAACTTCTCTGTTGGGAAATACGCAAGCTATCAT 1522
Db 1530 CCGTCTCTCAGGAACACCAAGAACTTCTCTCTGCGGGAAGCATGCAAGCTCTGCT 1589
QY 1523 GCAGAACTGATGAGAGATGAAAGTAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
Db 1590 GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649
QY 1583 GAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642
Db 1650 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1709
QY 1643 GTTCTGCTGATGAGACACATACGTTGATGAGTGTAGTGTGCTGCTGCTGCTGCTGCTGCT 1702
Db 1710 GCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
QY 1703 GAGCACAATTCAGAAAGACAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1762
Db 1770 GACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1829
QY 1763 GAGCATTGGAGTCAAGCAACCTTGGAGAGTGTGCGCTGCGGAGCTGTCTCAAGAGGA 1822
Db 1830 AAGCATTGGAATCAGACAGCACTTGAAGAGGTGCTGCTGCGGAGCTGTCTGGAAGCAGA 1889
QY 1823 GGTCAAGCATCACAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882
Db 1850 GGTCAAGCATCACAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949
QY 1883 CAAGCCAAAGCGCTGCGGCGCTTGTGAAATGATGATGATGATGATGATGATGATGATGAT 1942
Db 1950 CAAGCTGACGGCTGCGGCGCTTGTGAAATGATGATGATGATGATGATGATGATGATGAT 2009
QY 1943 GGGCAGAAAGAGCAGGCCCGCATTTTCAACCCAGCTCTCAAGACTCTCTTTCAGATGCT 2002
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on: February 1, 2004, 21:57:16 ; Search time 189 Seconds  
(without alignments)  
8164.420 Million cell updates/sec

le: US-09-042-460-1  
fect score: 3496  
uence: 1 GAATTCGGGTGGAGGCC.....CCGAGTCGGTACCAAGCTT 3496

ring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

ched: 569978 seqs, 220691566 residues

al number of hits satisfying chosen parameters: 1139956

inum DB seq length: 0  
inum DB seq length: 2000000000

t-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : Issued Patents NA:  
1: /cgm2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgm2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgm2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgm2\_6/ptodata/1/ina/PCUTUS\_COMB.seq.\*  
6: /cgm2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1600.8	45.8	4015	3	US-08-851-843A-224
2	1600.8	45.8	4015	3	US-08-974-549A-1
3	1600.8	45.8	4015	3	US-08-854-050-224
4	1600.8	45.8	4015	3	US-09-430-323-224
5	1600.8	45.8	4015	4	US-09-572-423B-3
6	1600.8	45.8	4015	4	US-09-128-354-1
7	1600.8	45.8	4015	4	US-09-675-321-1
8	1600.8	45.8	4015	4	US-09-052-919-1
9	1600.8	45.8	4015	4	US-08-912-951-1
10	1600.8	45.8	4015	4	US-09-733-294A-3
11	1596	45.7	4037	3	US-08-974-549A-343
12	1520.2	43.5	4029	3	US-08-851-843A-173
13	1520.2	43.5	4029	3	US-08-974-549A-232
14	1520.2	43.5	4029	3	US-08-854-050-173
15	1520.2	43.5	4029	4	US-09-430-323-173
16	1323.2	37.8	3855	3	US-08-974-549A-4
17	1323.2	37.8	3855	4	US-08-912-951-4
18	1138.2	32.6	3396	3	US-08-974-549A-638
19	1116.8	31.9	3451	3	US-08-974-549A-721
20	1112.4	31.8	3396	3	US-08-974-549A-639
21	939	26.9	3396	3	US-08-974-549A-642
22	896.4	25.6	3396	3	US-08-974-549A-641
23	786.6	22.5	3396	3	US-08-974-549A-640
24	716	20.5	2176	3	US-08-974-549A-3
25	716	20.5	2176	4	US-08-912-951-3
26	689.8	19.7	2171	3	US-08-851-843A-100
27	689.8	19.7	2171	3	US-08-974-549A-266

Sequence 100, App  
Sequence 100, App  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 30, Appli  
Sequence 62, Appli  
Sequence 8, Appli  
Sequence 62, Appli  
Sequence 62, Appli  
Sequence 31, Appli  
Sequence 9, Appli  
Sequence 213, App  
Sequence 139, App  
Sequence 48, Appli  
Sequence 60, Appli

ALIGNMENTS

RESULT 1  
US-08-851-843A-224  
; Sequence 224, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6093809el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/851,843A  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION:  
; APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:

## SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

/product= "hprt"

/note= "human telomerase reverse

OTHER INFORMATION: transcripase (hprt) catalytic protein

OTHER INFORMATION: component"

08-851-843A-224

very Match 45.8%; Score 1600.8; DB 3; Length 4015;

est Local Similarity 68.9%; Pred. No. 0;

atches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGGCCTTGAGCAATGACCCGCGCTCCTCGTTGCCCCGCGTGGCTCTCTGTGCG 82

40 CCCGGCACCCCGGATGCGCGGCTCCCGCTGCCGAGCCGTGCGCTCCCTGCTGG 99

83 CAGCCGATACCGGAGGTGCGCGCTGGCAACTTTTGTGCGCGCGCTGGGCGCCAGGG 142

100 CAGCCACTACCGGAGGTGCTGCGCTGGCAAGTTGCTGCGCGCGCTGGGCGCCAGGG 159

143 CAGCGCGCTTGCAACCCGGGACCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCC 202

160 CTGCGCGCTGTGACGCGCGGGACCCGCGGCTTCCGCGCGCTGCTGCGCGCACTG 219

203 AGTGTGATGCACTGGGCGCTCAGAGCTTCAAGCTTCAAGCTTTCCTTCAACAGGTGTC 262

220 GGTGTGCGTCCCTGGGACGACAGCGCGCGCCCGCGCGCTCCCTTCCGCGAGGTGTC 279

263 ATCCCTGAAGAGCTGTGGCGCGGTTGTGACAGACTCTGCGAGCGCAACAGAGAAA 322

280 CTGCTGAAGAGCTGTGTGGCCCGAGTGTGCAAGGCTGTGACGCGCGCGCGAGAA 339

323 CGTGTGCTTTTGGCTTTGAGCTGTCTTAAAGAGGCGAGGCGGCGCTTCCCATGCGCTT 382

340 CGTGTGCGCTTGGCTTGGCGCTGTGGAAGCGGCGCGCGCGCGCGCGCGCGCGCTT 399

383 CACTAGTAGCGGTAGCTACTTGGCCCAACATGTTATTGAGACCTTCGGTCTAGTGG 442

400 CACCAACAGCTGCGCACTACTTGCACCAACAGGTGACCGCACTGCGGAGGAGCGG 459

443 TGCAATGATGCTACTTGTAGCCAGTGGCGACGACCTGCTGTCTACTGCTGGCACA 502

460 GCGGTGGGGCTGTGCTGCGCGCGTGGGCGACGAGTGTCTGTTCACTGCTGSCAG 519

503 CTGTGCTCTTTATCTTCTGTGTGCGCGCGAGTGTGCTTACAGGTGTGTGGTCTCCCT 562

520 CTGCGCGCTCTTGTGTGTGTGTGCTTCCAGCTGCGCGCTTACAGGTGTGTGGGCGCGCT 579

563 GTACCAAAATTTGTGCCACAGGATATCTGCGCTCTGTGTGCGCTAGTTTACAGGCGCCAC 622

580 GTACAGCTCGCGGTGCTCATCTAGGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 639

623 CCGACCGGTGGGACAGAAATTTCACTAACTTTAGTTTCTTACAGATCAAGAGCAGTAG 682

640 GCGTCTGGGATGC-----GAACGGCGCTTGGAAACCATAG 672

683 TCGCCAGAGACACCGAAACCCCTGGCTTCCCATCTCGAGGTACAAAGAGGCATCTGAG 742

673 CGTCAAGAGGCGCGGGTCCCTCTGGGCTGCCAGCCCGGGTTCGAGGAGGCGCGGGG 732

743 TCTCAACAGTACAGTGTGCTTCACTTAAAGAGCCAGATGCTATCTCTCCCGAGAT 802

733 CAGTGCACCGCAAGTCTGCGGTGCGCGTGGCCCAAGAGGCGCGCGCTGCCCTTGAGCC 792

803 GGAGAGGACCCACAGCAGGTGCTACCAACCCCATCAGGCAATCATGTGGTCCAG 862

793 GGAGCGGACGCGCGTTGGGCAAGGGTCTCTGGGCCACCCCGGGCAGGACCGGTGACCGAG 852

863 TCGTCTCGGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTGTCTCTTAA 913

853 TGACCGTGGTTTCTGTGTGGTGTACCTGCGAGACCGCGCCGAAAGAGCCACTCTTTTGA 912

914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTGCGGTGTGCTGTAAACACAAAGCCAG 970

913 GGGTGGCTCTCTGGCACGCGCCACTCCACCCTCCGTGGCGCCGAGCACACCGCGG 972

971 CTCACAT---CTCTGTGTGACACCCCGGCAAAATGCTTTTTCAGTCTAGGCCATTTAT 1027

973 CCCCCCATCCATCGCGGCCACACGCTCCCTGGGACACGCTTGTCCCGCGGTGACG 1032

1028 TGAGACGACGATTTCTTTTACTCCAGGGAGATGGCCAAAGAGGCTTAAACCCCTCAT 1087

1033 CGAGACCAAGCACTTCTTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCT 1089

1088 CTTACTCAGCAACCTCCAGCCTTAACTTGACTTGGGCGCAGGAGACTGTGTGGAGATCAT 1147

1090 CTTACTCAGCTCTCTGAGGCCAGGCTGACTGGCGCTCGAGGCTCTGTGGAGACCATCT 1149

1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGCAAGGACACACCGCTCTATCGCGTC 1207

1150 TCTGGGTTCAAGGCTTGTATGCTCAGGAGCTCCCGCAGCTTGGCCGCTTCCCGCCAG 1209

1208 ATACTGGCAGATCGCGCCCTCTTCCAAAGCTGTGTGTAACCAATGAGAGTGCATA 1267

1210 CTACTGGCAATGCGGCCCTCTTCTGTGAGCTGCTTGGGAAACACACGCGCAGTGCCTTA 1269

1268 TGTCAAGCTCTCAGTCTACATTTGAGGTTTCGAACAGCAAAACCAAGGTGACAGATGC 1327

1270 CGGGGTGCTCTCTCAAGCAGCTGCGCGCTGGAGCTGGGTCAACCCAGCAGCGGTGT 1329

1328 CT-----TGAAACACAGGCC 1342

1330 CTGTGCGGGAGAACCCACAGGCTCTGTGGCGCCCGAGGAGGAGGACACAGACCC 1389

1343 ACCGACCTACTGAGTTTCTCGGCTGACAGCACTCCCTGGCAGGTATATGTTTCT 1402

1390 CCTCGCTGTTGACGCTCTCGCGACGACAGGCGCTTGGCAGGTGTACGGCTTCT 1449

1403 TCGGCGCTCTCTGCAAGGTGTGTCTGTAGTCTCTGGGTTACCGGACCAATGAGCG 1462

1450 GCGGCGCTGCTGCGCGCTGTTGCCCGCCAGGCTCTGGGCTCCAGGCAACAGACG 1509

1463 CCGTCTCTTTAAGAACTTAAAGATTTCTGTTGGGAAATACGGCAAGCTATCACT 1522

1510 CCGTCTCTCAGGAAACCAAGAAATTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1569

1523 GCAGGAAGTGTGGAAGATGAAGTAGAGATGCGCACTGGCTCCGAGCAGCGCCGCG 1582

1570 GCAAGGAGTCACTGGAAGATGAGCTGCGGACCTGCGCTGGCTGCGCAGAGGCCAGG 1629

1583 GAAGGACCGGTCTCCCGCTGCAAGACCGCTCTGAGGAGAGGATCTCTGGCTACGTTCT 1642

1630 GGTGGTGTGTTTCGGCGCCAGACCGCTCTGGCTGAGGAGATCTCTGGCCAGTTCT 1689

1643 GTTCTGCTGATGGAACACATAGCTGTACAGCTGTAGTCTATCTTTTACATCACA 1702

1690 GCACTGCTGATGAGTGTGCTGCTGAGCTGTCTCAGTCTTTCTTTTATGTCACGGA 1749

1703 GAGCACAATCCAGAGAACAGGCTCTTTCTTACCGTAAGAGTGTGTGAGCAAGCTCA 1762

1750 GACCACGTTTCAAAAGAACAGGCTCTTTTCTTACCGAAGAGTGTCTGGAGCAAGTGA 1809

1763 GAGCATTTGGAGTCAAGCAACCTTGAAGAGTGGGCTACCGGAGCTGTCTACAGAGA 1822

1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTTGAGCTGCGGAGGCTGTGGAAGACA 1869

1823 GGTCAAGCATCACAGACACCTGCTAGCCATGCCCATCTGCAAGACTGCGCTTTCATCCC 1882

1870 GGTCAAGCATTCGGGAAGGACGAGCCCGCTGCTGAGCTCCAGACTCCGCTTTCATCCC 1929

2963	QY	TCACGGCTCATTTCTAGACATTGACAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATA	3022
3001	DB		
3060	DB	TCACAGCGCTGTTTCTGATTTGACAGGTGAACAGCCTCCAGACGGTGTSCACCAACATCTA	3060
3023	QY	CAACATCTTCTGTTCTCAGGCCTACAGGTTTCATGCATGTGTGATTACAGTTCTCCCTTTGA	3082
3061	DB		
3083	QY	CAAGATCCTCTGCTGTCAGGCGTACAGGTTTTCACGCATGTGTGCTGCAGCTCCCATTTCA	3120
3083	QY	CCAGCGTTTGAAGAAAGCTCACATTTCTTTCTGGGCATCATCTCCAGCCAAAGCATCTGT	3142
3121	DB	TCAGCAAGTTTGAAGAAAGCCACATTTTCTTGCGCGTCATCTCTGACACGGCCTCCCT	3180
3143	QY	CTGCTATGCTATCTCTGAAGTCAAGAAATCCAGGAATGACATTAAAG-----GCTC	3193
3181	DB	CTGTACTCCATCTTGAAGCCAAAGAACCGAGGATGTGCTGGGGCCAAAGGGGCGCG	3240
3194	QY	TGGTCTCTTTTCTCTGAAGCCGCACATTTGGCTCTGTTACCAAGGCTTCTCTGCTCAAGCT	3253
3241	DB	CGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGGTGTGCCACCAAGCATTCTGCTCAAGCT	3300
3284	QY	GGCTGCTCATTTCTGTGCTATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT	3313
3301	DB	GACTCGACACGGTGTCACTTACGTGCCATCTCTGGGGTCACTCAGGACAGCCCGAGACGCA	3360
3314	QY	GCTTGGCGGAAGCTCCAGAGCGCACAAATGACCATTCTTAAAGTGCAGCTGACCCGAGC	3373
3361	DB	GCTGAGTCGAGAGCTCCGGGACGACGCTGACTGCCCTTGGAGGCCGAGCAACCCGGC	3420
3374	QY	CCTAAGCACAGACTTTCAGACCAATTTTGACTAA	3407
3421	DB	ACTGGCTTCAGACTTCAAGACCACTCTCGACTGA	3454

```

RESULT 2
US-08-974-549A-1
; Sequence 1, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

```

APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA: US 08/854,050  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA: US 08/911,312  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: US 08/912,951  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: US 08/915,503  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: WO PCT/US97/17618  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA: WO PCT/US97/17885  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTRT"  
OTHER INFORMATION: /note= "human telomerase reverse  
OTHER INFORMATION: transcriptase (hTRT) catalytic protein  
OTHER INFORMATION: component"  
-08-974-549A-1  
Query Match 45.8%; Score 1600.8; DB 3; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGGCGCTTGAGCAATGACCGCGCTCTCTGTTGCCCGCGGTGGCTCTCTGCTGG 82  
40 CCGGCGCAACCCCGATGCGCGGCTCCCGCTGCCGCGGTGGCTCTCTGCTGG 99  
83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGCGCCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGCGCCCGAGGG 159  
143 CAGGCGCTTGTGACCGGAGGCGGAGATCTACCGCACTTTGTGCGCAATGCCCT 202  
160 CTGGCGGCTGTGACCGGAGGCGGAGGCTTTCCGCGCTGTGCGCGCGCTGTGCG 219  
203 AGTGTGATGCTGAGGCTGACAGCTTCCAGCTTCCAGCTTCTTTCACCAAGGCTC 262  
220 GGTGTGCTGCTGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
263 ATCCCTGAAGAGCTGGTGGCGAGGTTGTGAGAGCTCTCGAGGCGCAACGAGAAA 322  
280 CTGCTGGAAGAGCTGGTGGCGAGGTTGTGAGAGGCTGTGAGAGGCGCGCGAGAA 339  
323 CGTGTGCTGCTTTGGCTTTGAGTGTGTTAAACGAGGCGAGAGCGCGGCTTCCCATGGCCTT 382  
340 CGTGTGCTGCTTTGGCTTTGAGTGTGTTAAACGAGGCGAGAGCGCGGCTTCCCATGGCCTT 399  
383 CACTAGTAGGCTGCTAGTACTTGGCGCGAGCTGTTATTGAGACCTGCTGCTAGTGG 442

400 CACCACCGGTGGCGCACTACCTGCTGCCAACACACGGTGCACCGACCTGCGGGGAGCGG 459  
443 TGCATGAGTACTGTTGAGCGAGTGGGCGACGACTGCTGCTTACCTGCTGGCACA 502  
460 GCGGTGGGGGTGCTGCTGCTGCGCGGTGGGCGACGAGTGTGCTTACCTGCTGGCAG 519  
503 CTGTGCTCTTTATCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 562  
520 CTGCGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAAAATTTGTGCCACACGAGATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639  
623 CCGACCGGTGGGCGAGAAATTTCACTTAACCTTAGTCTTCAACAGATCAAGAGCAGTAG 682  
640 GGTCTGGGATGC-----GACGGGCTGGAACCATAG 672  
683 TCGCCAGGAGCACCGAAACCCCTGCGCTTGCATCTCGAGGTACAAAGAGGATCTGAG 742  
673 CGTCAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732  
743 TCTCACCAGTCAAGTGTGCTTCAAGTAAAGAGCGCAGATGCTATCTGCTCCGAGAGT 802  
733 CAGTCCAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
803 GAGGAGGAGCACCGCAGAGTGTCTACCAACCCCATCAAGCAAAATCATGGGTGCCAAG 862  
793 GAGCGGAGCGCGCTTGGGCGAGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 852  
863 TCTGCTCGTCCCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
914 AGAAAGGTGCTGCACTGAGTCTCTC---TGGGTGGGTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGTGCGCTCTGCGCACGCGCTCTCCACCATCGTGGCGCGCGCGCGCGCGCGCGCG 972  
971 CTCACAT---CTCTGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1027  
973 CCGCGCATCCACATCG 1032  
1028 TGAGACCAAGCATTTCTTTTACTTCCAGGGGAGATGGCGAAGCGCTTAAACCCCTCAT 1087  
1033 CGAGACCAAGCATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089  
1088 CTACTCAGCAACTCAGCCTTAACTTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1147  
1090 CTTACTAGTCTCTGAGCG 1149  
1148 TCTGGCTCAAGGCTTAGGACATCAGGACCATCTCTGCGAGCACACCGCTTATCGCGTGG 1207  
1150 TCTGGTTCAGGCG 1209  
1208 ATATGCGAGATGCG 1267  
1210 CTACTGCAAAATGCG 1269  
1268 TGTCAAGCTCTCAGGTCACTTTCAGAGTTTCGAAACAGCAAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTTCAAGACGCACTGCGCGCTGCGAGTGGGTGCGGTGCGGTGCGGTGCG 1329  
1328 CT-----TGACACACGCGCG 1342  
1330 CTGTGCGCGGAGAACCG 1389  
1343 ACCGACCTCATGATTTGCTCGCGCTGACAGCGTCCCTGCGAGGTATATGTTTCTTCT 1402  
1390 CCGTGCCTGCTGAGTGTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1449  
1403 TCGGCGCTGCTCTGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1462  
1450 GCGGCGCTGCTGCG 1509

1463 CGCTTCTTTAAGAACTTAAGAACTTCACTCTGTTGGGAATACGGCAAGCTATCACT 1522  
1510 CGCTTCTCTCAGGAACCAAGAACTTCACTCTGTTGGGAATACGGCAAGCTATCACT 1569  
1523 GAGGAACTGATGTGGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1582  
1570 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1629  
1583 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1642  
1630 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1689  
1643 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1702  
1690 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1749  
1703 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1762  
1750 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1809  
1763 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1822  
1810 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1869  
1823 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1882  
1870 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1929  
1883 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1942  
1930 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 1989  
1943 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2002  
1950 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2049  
2003 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2062  
2050 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2109  
2063 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2122  
2110 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2169  
2123 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2182  
2170 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2229  
2183 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2242  
2230 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2286  
2243 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2302  
2287 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2346  
2303 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2362  
2347 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2400  
2363 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2422  
2401 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2460  
2423 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2482  
2451 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2520  
2483 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2542  
2521 GAGGAGCTGAGTGAAGTGAAGTAGAGATGCACTGGCTCGGAGAGAGATCTGCTTCT 2580

QY 2543 CAGTCTGTTTGGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGATGAGAGAT 2602  
Db 2581 CAGCTGTTGTTACGGGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGAT 2640  
QY 2603 GCTTTTACGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2662  
Db 2641 GCTCTCGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
QY 2663 CTTCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
Db 2701 CTTCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
QY 2723 GACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782  
Db 2761 GACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
QY 2783 GCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
Db 2821 GCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
QY 2843 GTTCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2902  
Db 2881 GCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
QY 2903 TGTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
Db 2941 CGGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
QY 2963 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3022  
Db 3001 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060  
QY 3023 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082  
Db 3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
QY 3083 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
Db 3121 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
QY 3143 CTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3193  
Db 3181 CTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240  
QY 3194 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
Db 3241 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
QY 3254 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3313  
Db 3301 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
QY 3314 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3373  
Db 3361 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
QY 3374 CTTAAGCAGAGCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTC 3407  
Db 3421 ACTGCTCAGAGCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTCAGAGCTTTC 3454

## RESULT 3

US-08-854-050-224  
; Sequence 224, Application US/08854050  
; Patent No. 6261836  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.



TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536

## ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-0029303US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
OTHER INFORMATION: /note= "human telomerase reverse  
OTHER INFORMATION: transcriptase (hTERT) catalytic protein  
OTHER INFORMATION: component"

-08-854-050-224

Query Match 45.8%; Score 1600.8; DB 3; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGGACACATGACCCGCTCTCTGTTGCCCGCGGTGGCGCTCTCTGCTGCG 82  
40 CCGGCGCACCCCGGATGCGCGCGCTCTCCGCTCCGAGCCGTGCGCTCCCTGCTGCG 99  
83 CAGCCGATACCGGAGGTGGCGGTGCGCAACTTTGTGCGGCGCTTGGGCGCCGAGGG 142  
100 CAGCACTACCGGAGGTGCTGCGCTGCGCACTTGTGCGGCGCTTGGGCGCCGAGGG 159  
143 CAGGCGCTTGTCCACCCCGGACCGGAGATCTACCGCACTTTGGTGTGCCAATGCT 202  
160 CTGGCGGCTGGTGGCGGCGGACCCGCGGCTTTCCGCGGCTGGTGGCGCCGCTGCT 219

QY 203 AGTGTGATGCACTGGGGCTCAGAGCTCCAGCTCGCGACCTTTCTCTCCACCGAGTGC 262  
DB 220 GGTGTGGTGTCTGGAGCGCAGCGCGCGCCCGCCCTCTCTCCGCGAGTGC 279  
QY 263 ATCCCTGAAAGAGCTGTGGCCAGGGTTGTGAGAGACTTCTCGAGCGCAACGAGAGAA 322  
DB 280 CTGCTGAAGAGCTGTGGCCGAGTGTGAGAGGCTGTGAGAGCGCGCGCGGAGAA 339  
QY 323 CGTCTGGCTTTGGCTTTGAGTGTAAAGAGCGCAGAGCGGCGCTCCCATGGCTTT 382  
DB 340 CGTCTGGCTTTGGCTTTGAGTGTAAAGAGCGCAGAGCGGCGCTCCCATGGCTTT 399  
QY 383 CACTAGTAGCTGTGCTAGCTACTTGGCCCACTGTATTGAGACCTCTGCTGTGAGTGG 442  
DB 400 CACCAACAGCGTGGCAGCTACTCTGCCCAACAGCGTACCGACGCACTCGGGGAGCG 459  
QY 443 TGCAATGATGCTACTTGTAGCGAGTGGCGAGACCTGTGCTGTCTACTCTGTGCA 502  
DB 460 GCGTGGGGGCTGCTGTGCGCGCGTGGCGGAGCGAGTGTGCTGTGCTGTGCA 519  
QY 503 CTGTGCTTTTATCTTCTGTGCTGCGCCAGCTGTGCTTACCAGCTGTGCGGTCTCCCT 562  
DB 520 CTGCGCGCTCTTGTGTGTGTGCTCCAGCTGCGGCTACCAGTGTGCGGCGCGCT 579  
QY 563 GTACCAATTTGTGCCACACCGATATCTGCGCTCTGTGCTGCTGTGCTGTGCTGTG 622  
DB 580 GTACCACTGCGGCTGCTCACTCAGGCGCGCGCGCCGCGCACACGCTAGTGGACCC 639  
QY 623 CCGACCGCTGGCGAGGAAATTCATCACTTAGTCTTACACAGATCAAGAGAGTAG 682  
DB 640 GCGTCTGGATGC-----GAAAGCGCTTGNACCATAG 672  
QY 683 TCGCCAGGAAGCACCGAAACCCCTTGGCTTGCATCTCGAGGTACAAAGAGGATCTGAG 742  
DB 673 CGTCAAGGAGCGCGGCTCCCTGCGCTGCCAGCGCGCTGCGAGGAGCGCGGCG 732  
QY 743 TCTACAGTACAGTGTGCTTCCAGTACAGAGCGCAGATCTATCTCTGCTCCGAGAGT 802  
DB 733 CAGTGTCCAGCGGAGTGTGCGTGTGCCAGAGAGCGCGCGCTGCGGCTGCGCTGAGCC 792  
QY 803 GGAGGAGGAGCGCGCAGGAGTGTACCAACCCCATCAGGCAAAATCATGGTGTGCAAG 862  
DB 793 GGAGCGGAGCGCGCTTGGCAGGAGTCTTGGCGCGCGCGCGCGGAGCGGTGAGCCGAG 852  
QY 863 TCTGCTCGTCTCCCGGAGGTGCTACT-----GCAGAGAGATTTGTCTTAA 913  
DB 853 TGACCGTGGTTTCTGA 912  
QY 914 AGGAAAGGTGTGTGACCTGAGTCTCTC---TGGGTGCGTGTGTGTGTGTGTGTGTGTGT 970  
DB 913 GGTGCGGTCTTGTGCGCGCGCTTCCAGCGCGCTTCCAGCGCGCGCGCGCGCGCGCG 972  
QY 971 CTCCACAT---CTGT 1027  
DB 973 CCCCCATCCACATCGCGGCGCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032  
QY 1028 TGAGACCGACATTTCTTACTTCCAGGAGAGTGGCAAGAGCGTCTAAACCCCTCAT 1087  
DB 1033 CGAGACCGAGACATTTCTTACTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1089  
QY 1088 CCTACTGAGCAACCTTCCAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147  
DB 1090 CTTACTGAGTCTCTGTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
QY 1148 TGTGGCTCAGGCGCTAGGACATCAGGACCTCTGCGAGGACACACCGCTCTATCGCGTGC 1207  
DB 1150 TGTGGTTCAGGCGCGCTGATGCGAGGAGCTTCCCGGAGGTGCGCGCGCTTCCCGAGCG 1209  
QY 1208 ATACTGGCAGATGCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1267  
DB 1210 CTACTGGCAATGCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269  
QY 1268 TGTGAGCTCTCAGGTCAATTTGAGGTTTTCGAGCGAGCAACCAACAGGTGTGAGATGC 1327

1270 CGGGGTGCTCTCAAGAGCGCACTGCCGCTGCGAGTGGCGGTCACCCCAGACGCGGTGT 1329

1328 CT-----TGAACACACAGCCC 1342

1330 CTGTGCCGGGAGAACCCCAAGGCTCTGTGGCGCCCCCGAGGAGGAGACACAGACCC 1389

1343 ACCGCACCTCATGGATTTGCTCCGCTGCGACAGAGTCCCTGGCAGGTATATGTGTTTTCT 1402

1390 CCGTCGCCCTGGTGACGTGCTCCGCCAGACAGACAGCCCTGGCAGGTGTACGGCTTCGT 1449

1403 TCGGGGCTGTCTCTCAAGGTGGTGTCTGTAGTCTCTGGGGTACACAGGCAATAGAGC 1462

1450 GCGGGCTCCTCGCGCGCTGTGCCCCAGGCTCTGGGGCTCCAGGCACACAGAACG 1509

1463 CCGCTCTTTAAGAACTTAAGAAGTTCACTCTGTTGGGGAATACGGGAAGCTATCACT 1522

1510 CCGCTTCCTCAGGAACACCAAGAAGTTCACTCTCCCTGGGGAAGCATGCGCAAGCTCTCGCT 1569

1523 GCAGGAACGTAGTGTGGAAGATGAAGTAGAGGATGSCACTCGCTCCGACAGCCCGG 1582

1570 GCAGGAGCTGAAGTGTGAAGATGAGCTGCGGGACTGGGCTTGGCTGGCAGAGCCGAGG 1629

1583 GAAGGACCTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCCTGGGTAGCTTCCT 1642

1630 GGTGTGCTGTGTCCGGCGCAGACACCGCTCTGCGTGAGGAGATCCTGGCAAGTTCTCT 1689

1643 GTTCTCGCTGATGGACACATAGTGTACAGCTGCTTAGGTCACTTTTATCATCACAGA 1702

1690 GCACTGGCTGATGAGTGTGACGTGCTGAGAGTGTCTCAGGTCTTCTTTTATGTACGGA 1749

1703 GAGCACATTCAGAGAAACAGGCTCTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762

1750 GACCACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGCAAGTTGCA 1809

1763 GAGCATGTGAGTCAGGCAACACTTGAGAGAGTGGGCTACGGAGCTGTCAAGAGGA 1822

1810 AAGCATTGAATTCAGACAGACTTGAAGAGGGTGCAGCTGCGGAGCTGTGGAAGCAGA 1869

1823 GGTCAAGCATCACAGGACACCTGGCTAGCCATGCCCATCTGCAGACTGCCCTTCATCCC 1882

1870 GGTCAAGCAGCATCGGAAGCAGCGCCGCTGTGACGTCCAGCTCCGCTTCATCCC 1929

1883 CAGCCCAAGCGCTCGGCGCATTTGGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942

1930 CAAGCCTGACGGGCTCGGCGCGATTTGTGAACATGGACTAGCTCGTGGGAGCGAGAGCTT 1989

1943 GGCAGAGAAGACAGGCGCCAGCAATTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002

1990 CCGCAGAGAAAGAGCGCGAGCTCTCACTCGAGGGTGAAGGCACTGTTCAAGCTGCT 2049

2003 CAACATGAGCGGAACAAACATCTCACTTATGGGGTCTTCTGACTGGGTATGAATGA 2062

2050 CAACTACGAGCGGGCGGCGCGCCGCGCTCTGGGCGCTCTGTGCTGGGCTGGGACGA 2109

2063 CATCTACGAGCTCGCGGGCTTTGTGCTCGCGTGGCGCTCTGGACCAAGACACCCAG 2122

2110 TATTCACAGGGCTGGCGACCTTGTGCTGCGGTGGGCGCGGAGACCCGCGCGCTGA 2169

2123 GATGTACTTTGTTAAGGCAGATGTGACCGGGGCTATGATGCGATCCCGCAGGCTAAGCT 2182

2170 GCTGTACTTTGTCAAGTGGATGTGACGGCGGTACGACACCATCCCGCAGGACAGCT 2229

2183 GGTGGAGGTTGTTGCCATATGATCAGGCACTCGGAGACAGCTACTGTATCCGCCAGTA 2242

2230 CAGCGAGGTCACTCGCCAGCATCACTCA--AACCCCGAGAACAGCTACTCGTGGCTCGGTA 2286

2243 TGCAGTGGTCCGAGAGATAGCAAGGCCAAGTCCACAAAGTCTTTAGGACAGAGGTCAAC 2302

2287 TGGGTGGTCCAGAGGCGGCCATGGGCACGTCCGAGAGGCTTCAAGAGCCACGTCTC 2346

2303 CACCTCTCTGACCTCGAGGCATACATGTGGGCGAGTTCTTTAAGCATCTGCAAGGATTCAGA 2362

2347	TACCTTGA	CAGACCTC	CAGCCGCT	ACATGCG	CAGATT	CGTGCGT	CACCTG	CAG-----GA	2400	
2363	TGCAGTGC	ACTCAGGAA	CTCCGTTG	TCATCG	AGCAGC	AGCATCT	CTATGA	ATGAGAGCAG	2422	
2401	GACACGCG	CTGAGGAT	GCCGTG	CTCATCG	AGCAGAG	CTCTCC	TGATG	AGGCCAG	2460	
2423	CAGAGCCT	GTGTTGAC	ATTCTT	CTGTG	CACATT	CTCTG	GTG	CTGTAAGATGTGTGA	2482	
2461	CAGTGGCT	CTTTGCAG	CTCTCT	CAAGCTT	CTATGTG	CCACAC	CGCGT	GCGCATCAGGGG	2520	
2483	CAGGTGCT	ATACG	CAGTGC	CAGGCG	CAATCCC	CAGAGG	CTCCAG	CCCTATCCACCTGCTCTG	2542	
2521	CAAGTCT	TAAGT	CAGTGC	AGGGA	TCCG	CAGAGG	CTCCAT	CTCTCCACGCTGCTCTG	2580	
2543	CAGTCTGT	TTTGGG	ACATATG	AGAA	CAAGCT	TTGTTG	CTAG	GTGTCAGCGGATGCGTT	2602	
2581	CAGCCTGT	GTCTAC	GGCACA	TATGAG	AA	CAAGCT	TTTGC	GGGATTTCCGCGGGACGCGCT	2640	
2603	GCATTTAC	GTGTTG	TATGAC	TTTTCT	GTGTTG	TGAG	CGCTCA	CTTTGGACCAAGCAAAAC	2662	
2641	GCTCTG	GTGTTGG	ATGATTT	CTTGT	TTGGT	AGAA	CACTCA	CTTACCCACGCGAAAC	2700	
2663	CTTCTC	AGCACCT	TGTCAT	CGCGT	TCTT	GAGTAT	GGGTG	GCATATAAACTTGCAGAA	2722	
2701	CTTCTC	CAGGAC	CCGTGTC	AGG	TGCTCC	TGATAT	GGCTG	CGTGGTGAACCTTGC	2760	
2723	GACAGT	GTGTA	CTTCCCT	GTGAG	CCTGGT	TACCTGG	TGTCAG	CTCCATCAGACT	2782	
2761	GACAGT	GTGTA	CTTCCCT	GTGAA	GAACAG	AGCCCTGG	TGGC	ACGGCTTTTGTTCAGAT	2820	
2783	GCCTGCT	CAC	TGCCCT	GTGTTCC	CTGTG	TGGCTT	GTGTG	GACACCTCAGACTTTTGAGAGT	2842	
2821	GCGGCCCC	ACGGCCTA	TTCCCT	CTGGT	CGGCTG	CTGCTGG	TATAC	CCCGGACCTGAGAGT	2880	
2843	GTTCGT	GACTACT	CAAGTT	ATGCC	CAGAC	CTCAAT	TAAG	ACAGAGCTCACCTTCCAGAG	2902	
2881	GCAGAG	CGACTACT	CTC	AGCTAT	CGCCGG	ACCTCCT	CAATCAG	AGCCAGCTCTCAGCTTCAACCG	2940	
2903	TGCTTT	CAAA	GCTCGG	AAAC	CAATGCGG	AA	CAAGCT	CTCTGTGGTCTTGCGGTTGGAAGTG	2962	
2941	CGGCTT	CAAG	GCTGG	AGGA	AA	CAATGCT	CGCA	AACTCTTTGGGGTCTTGCGGCTGAAGTG	3000	
2963	TACGGT	CTATTT	CTAG	CTTG	CAGT	GAA	AGCCT	CCAGACAGCTGTGCATCAATATA	3022	
3001	TCA	CAGCCT	GTGTTCT	GATTTG	CAGGT	GAAC	AGCCTCC	CAGACGGTGTGCACCAACATCTA	3060	
3023	CAAGAT	CTTCTG	CTCAG	GCCTT	CACAG	TTTCA	TGATGTG	TATCAGGCTTCCCTTTGA	3082	
3061	CAAGAT	CTCTCT	GTG	CAGG	CTTAC	AGGTTT	CA	CGATGTGTCTGAGCTCCCATTTCA	3120	
3083	CCAGCG	TTTAGG	AAAC	CTCAC	ATTCTT	CTGGG	CA	TCATCCACGCAAGCATCCCTG	3142	
3121	TCAG	CAAGTT	TTGGA	AA	CCCCA	TTTTTCT	CTCG	CTCATCTGACCGGCTCCCT	3180	
3143	CTGCTAT	CTATCT	CTGAG	GTCA	AGATCC	AGGATG	CA	CTAAG-----GCCTC	3193	
3181	CTGCTACT	CTCCAT	CTGAA	AGCC	AAAGAA	CGAGG	ATGCGT	GTGGGGCCAAAGGCGCGCGC	3240	
3194	TGGCTCT	CTTTCT	CTGAG	CCGCA	CA	TTTGGCT	CTGTGCT	TACCAGGCTTCTCTGTCAAGCT	3253	
3241	CGG	CCCTCT	CTG	CCCTCG	AGG	CCGTG	CA	GTGGCTGTGCCCAAGCATTCCTGTCAAGCT	3300	
3254	GGCTGCT	CA	TTCTG	CA	TCTA	CAAA	TGTTCT	CTGGGA	CTCTGAGGACAGCCCAAAACT	3313
3301	GACTCG	ACA	CGGTG	CA	CTTCA	CGTGC	CA	CTCTCC	TGGGTCATCTCAGGACAGCCACGCA	3360
3314	GCTGTG	CGGAAG	CTCC	AGAG	CGCA	ATGAC	CACTCT	TAAAGCTG	CAGCTGACCCAGC	3373
3361	GCTGAG	TCGG	AA	GTCTCC	GGG	AG	AG	CGCTG	AGCGCCGACGCCACCGGC	3420
3374	CCTAAG	AC	AG	CACTTT	CAG	ACCA	TTT	GACTAA	3407	
3421	ACTGCC	CTCAG	ACTTCA	AGACCA	CTCT	CG	ACTG	CA	3454	

27

RESULT 4  
S-09-430-323-224  
Sequence 224, Application US/09430323  
Patent No. 6309867

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. 6309867el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/430,323

FILING DATE: 29-Oct-1999

CLASSIFICATION: &lt;Unknown&gt;

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 08-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:

## SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

## FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"

/note= "human telomerase reverse

transcriptase (hTRT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

JS-09-430-323-224

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

27

23 CCGGGCTTGGACAAATACCGGGCTCTCGTGGCCCGGGTGGCGCTCTGTGGCG 82

Db 40 CCGGGCCACCCCGCGATGCCGCGCTCCCGCTGCGAGCGGTGGCTCCCTGCTGG 99  
QY 83 CAGCCGATACCGGAGGTGTGGCCGCTGGCAACCTTTGTGCGGCGCTGGGGCCCGAGGG 142  
Db 100 CAGCCACTACCGGAGGTGTGGCCGCTGGCCACGTTGTGCGGCGCTGGGGCCCGAGGG 159  
QY 143 CAGCGGCTTGTGCAACCCCGGAGACCGGAAGATCTACCGCACTTTGGTTGCCCAATGCTT 202  
Db 160 CTGGCGCTGTGACGCGCGGAGACCCGGGGCTTTCCGGCGCTGGTGGCCAGTGGCT 219  
QY 203 AGTGTGATGACCTGGGCTCA CAGCCTCCACTGTCGCGACCTTTTCTCCACAGGTGTC 262  
Db 220 GGTGTGGCTGCTTGGGACGACGCGCGCGCCCGCCCGCGCTCTCTCCGCGAGGTGTC 279  
QY 263 ATCCCTGAAAGAGTGTGGCCAGGTTGTGCAGAGACTCTGCGAGCGCAACGAGAGAAA 322  
Db 280 CTGCTGAAGAGCTGTGGCCCGAGTGTGCAGAGGCTGTGCGAGCGCGGCGGAGAA 339  
QY 323 CGTCTGGCTTTTGGCTTTGAGCTGTAA CAGAGGCCAGAGCGGGCTTCCCATGGCTT 382  
Db 340 CGTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 399  
QY 383 CACTAGTAGGTGGTAGCTACTTGCACACACTGTTATTGAGACCTCGCTGTCAGTGG 442  
Db 400 CACCACAGCGTGGCAGCTACTTGCCTCCACACGCTGACCGGACCGGAGCGGAGCG 459  
QY 443 TGCATGATGCTACTTGTGACCGAGTGGGCGACGACCTGCTGTCTACCTGCTGCGACA 502  
Db 460 GGGTGGGGCTGCTGCTGGCGCGCTGGGCGGACGAGTGTGTTACCTGCTGCGACG 519  
QY 503 CTGTGCTTTTATCTTGTGTCGCCCGAGCTGTGCTACCAAGTGTGTGGTCTCCCT 562  
Db 520 CTGCGCGCTCTTGTGTGGTGGCTCCAGCTGCGCTTACAGGTGTGCGGGCGCGCT 579  
QY 563 GTACCAAAATTTGCCACACCGATATCTGGCCCTCTGTGCTCGCTAGTTACAGGCCAC 622  
Db 580 GTACAGCTCGGCGCTCCACTCAGGCCCGGCCCGCCACACGCTAGTGGAGCCCGAAG 639  
QY 623 CCGACCGGTGGGAGGAATTTCACTAACCTTAGTGTCTTACACAGATCAAGAGCATGAG 682  
Db 640 CGCTCTGGGATG-----GACCGGGCTGGAACCATAG 672  
QY 683 TCGCAGGAGACCCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
Db 673 GTGAGGAGCGCGGGTCTCCCTTGGCTTGCAGCCCGGGTGCAGAGAGCGCGGGGG 732  
QY 743 TCTCACCAGTACAAAGTGTGCTTCACTAAGAGGCGAGATGCTATCTGTCCCGAGAT 802  
Db 733 CAGTGCACAGCGAAGTCTGCGCTTGGCCAGAGGCGCGCTGGCGCTGCCCTGAGCC 792  
QY 803 GGAGGAGGACCCACAGGCGAGTGTACCAACCCCATCAGGCAATCATGGTGCCAAG 862  
Db 793 GGAGGAGCGCGCTTGGGAGGCGCTGGGCGCCACCGGCGGAGGACCGGTGGAGCCGAG 852  
QY 863 TCCTGCTCGGTCCCCGAGGTGCTACT-----GACAGAGAAAGATTTGCTTTAA 913  
Db 853 TGACCGTGGTTTCTGTGTGTGTCACTGCGACAGCCCGCGGAGAGCCACCTCTTGA 912  
QY 914 AGGAAAGTGTGACCTGAGTCTCTC---TGCGTGGTGTGCTGTAAACACAGCCAG 970  
Db 913 GGGTGGCTCTCTGGCAGCGGCCACTCCCAACCATCCGTGGGCGCGGCGGAGCACCGCGGG 972  
QY 971 CTCACAT---CTCTGTGTACCAACCCCGCCAAAATGCCCTTTTCAGCTCAGGCCATTTAT 1027  
Db 973 CCCCCATCCACATCGGGCCACCACTGCTCCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
QY 1028 TGAGACGACATTTCTCTTCTCAGGAGAGATGGCAAGAGCGCTTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCT 1089  
QY 1088 CCTACTCAGCAACCTCCAGCCTAACTTACTTGGGCGCAGGAGACTGTTGAGATCATCTT 1147  
Db 1090 CTTACTAGCTCTCTGAGGCCAGCTGACTGCGGTGCGAGGCTCTGTGAGACCATCTT 1149

1148 TCTGGGCTCAAGGCTTAGGACATCAGAGACACCTCTGCAGGACACACCGTCTATCAGCGTG 1207  
1150 TCTGGGTTCCAGGCGCTGTGATGACAGGAGCTCCCGCAGGTTGCGCGCTGCCCGCAGCG 1209  
1208 ATACTGGCAGATGGCGGCTCTGTTCCAAAGCTGCTGGTGAACCATGACAGAGTGCCAAATA 1267  
1210 CTACTGGAAATGGCGGCTCTGTTCTGAGCTGCTTGGAAACCAAGCGCAGTGCCCTTA 1269  
1268 TGTGAGACTCCTCAGGTCAATTGCAAGTTTGAACAGCAAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGCACTGCCGCTGCGAGCTGCGGTCAACCCAGCAGACCGGTGT 1329  
1328 CT-----TGAAACACCAAGGCC 1342  
1330 CTGTGCCGGAGAGCCCGCAGGCTCTGTGGCGGCCCGGAGGAGGAGACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGACACAGCTGCTCCAGGTATATGTTTCT 1402  
1390 CCGTGGCTGTGTCAGCTGCTCCGACAGCAGACGCTGCGAGGTACGGTTGCT 1449  
1403 TCGGGCTGTCTGCAAGGTGTGCTGTAGTCTCTGGGGTACAGGCAAAATGAGGG 1462  
1450 GCGGGCTGCTGCGCGCTGTGTGCCCCAGGCTCTGGGGCTCCAGGCAAAAGG 1509  
1463 CCGCTCTTTAAGAACTTAAAGAACTTCACTCTGTTGGGGAATAGCGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAACTTCACTCTCTGGGGAAGCATGCCAAGCTCTCGT 1569  
1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATGGCACTGGCTCCGACAGCGCCGG 1582  
1570 GCAGGAGTGAAGTGAAGATGAGCGTGGCGGACTCGCTTGGCTGCGAGAGGCCAGG 1629  
1583 GAAGACCGTGTCCCGCTGCAGAGCACCTGTGAGGAGAGGATCCTGGCTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGCGCAGAGCACCTGTGCTGAGGAGATCTGGGCCAAGTTCT 1689  
1643 GTTCTGGCTGATGACACATACGTTGTPACAGCTGTAGTCACTTCTTTTATCATCAGAA 1702  
1690 GCACTGGCTGATGATGTGATGCTGCTGAGCTGCTCAGGCTCTTTCTTTTATGTCAGGA 1749  
1703 GAGCACATTCAGAGAACAGGCTCTTCTTCTACCTAGAGTGCTGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACACCTTGAGAGATGCGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTGCGAAGCAGA 1869  
1823 GGTGAGGATCACAGACACCTGGCTAGCCATGCCATCTGCAGACTGCGTTTCATCCC 1882  
1870 GGTGAGGAGCATCGGAAGCCAGGCGCGGCTGCTGACGTCAGACTCCGTTTCATCCC 1929  
1883 CAAGCCAAACCGGCTCGCGGCCATTGTGAACATGAGTTATAGCATGGGTACCAAGCTTT 1942  
1930 CAAGCTCAAGGCTGCGGCGGATTGTGAATGAGTACGTCGTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGCAGCGCCAGCATTTTCACCAAGCTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACCTGAGGGTGAAGGCACTGTTTCAGGGTGT 2049  
2003 CAATATGAGCGGCAAAACATCTCACCTTATGGGCTCTTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGGCGCGGCGCCCGGCTCTCTGGGGCTCTGTGCTGGGCTGGAAGA 2109  
2063 CATCTACAGGACTGGCGGCGCTTGTGCTGCGTGTGCGTGTCTGGACCAAGCAAGCCAG 2122  
2110 TATCCAGAGGCGCTGGCGCACTTCGTGCTGCGTGTGCGGCGCCAGAACCCCGCGCTGA 2169  
2123 GATCTACTTTTAAAGCATGATGACCGGGGCTATGATCCATCCCGCAGGTAAGCT 2182  
2170 GCTGTACTTTTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCCAGGAGCGGT 2229

2183 GGTGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGGAGTTCATCGCCAGCATCATCA---AACCAGAAACAGTACTGCGTGTGCGTA 2286  
2243 TGCAGTGGTCCCGAGAGATAGCAAAGGCAAGTCCCAAGTCCCTTTAGGAGACAGGTCCAC 2302  
2287 TGCCTGGTCCAGAAAGCGGCCCTTGGCACAGTCCGCAAGGCCCTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAAGCATCTGCAGGATTCAGA 2362  
2347 TACCTTTGACAGACCTCCAGCCGTATACGCAAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCACCTGAGGAACCTCCGTTGTCTCATGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCACCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGAATCTTCTCTGCACTTCTCTGGTCAAGTGTCTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTTTCGACGCTCTCTACGCTTCATGTGCCACCACCGCTGCGCATCAGGGG 2520  
2483 CAGTCTATACGCAAGTGCAGGCACTCCCGAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAAGTCTCTAGTCCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTGGAGACATGAGAAAGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGACATGAGAAAGCTGTTTGGGGGATTCGCGGGGACGGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTGGTGACCTCCTCTGAGCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTTGTGGTGACCTCACCTCACCCACGCGAAAC 2700  
2663 CTTCTCTCAGCACCTTGGTCCATGGGCTTCTGAGTATGGGTGCATGATAAACTTGCAGAA 2722  
2701 CTTCTCTCAGGACCTTGTCTCGAGGTGTCCTGAGTATGCTGCGTGGTGAACCTTGGCGAA 2760  
2723 GACAGTGTGAACCTTCCCTGTGAGCCTGTATCCCTGGTGGTGGAGCTCCATACCAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTAGAAAGCAGGAGCCCTGGGTGGCAGCGCTTTTGTTCAGAT 2820  
2783 GCCTGTCTACTGCTGTTTCCCTGCTGCTGCTGCTGCTGACACTCAGACTTTTGGAGGT 2842  
2821 GCGGGCCACAGGCTATTCCTCTGTGCGGCTGTCTGCTGATACCGGACCCCTGGAGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAAGACGAGCTTCACTTCCAGAG 2902  
2881 GCAGAGCACTACTCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAGAACCATGCGGAAAGCTCCTGCTGCGTCTTTCGGTGTGAAGTG 2962  
2941 CGGCTTCAAAGCTGGGAGGAACATGCTGCGAAACTCTTTCGGGCTTTCGGCTGAAGTG 3000  
2963 TCAGGCTCTATTCTAGACTTCAGGTGAACAGCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCTCCAGACCGTGTGCACCAACATCTA 3060  
3023 CAGATCTTCTGCTTCAGGCTTACAGTTCCATGCTGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTCTCTGCTGACGGCTACAGGTTTCAACGATGTGTGCTGAGCTCCCATTTCA 3120  
3083 CAGGCTGTGTAGGAAGAACCTCACATTTCTTCTGGGCATCATCTCCAGCCAGCATCCTG 3142  
3121 TCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGCTCATCTCTGACACGGCTCCCT 3180  
3143 CTGCTATGCTATCTGAGGTCAAGATCCAGGATGACATAAG-----GCTC 3193  
3181 CTGCTACTCATCTCTGAAGCCAGAACGAGGATGTGCTGGGGGCCAAGGGCGCGCG 3240  
3194 TGCTCTCTTCTCTGAAGCCGACATTTGGCTCTGCTACAGGCTTCTCTGCTCAAGCT 3253  
3241 CGGCTCTCTGCTCTCGAGGCGCTGAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGCT 3300  
3254 GGCTGTCTATCTGTCTCATCTACAAATGTCTCTCTGGGAGCTCTGTGAGGACAGGCCAAAACCT 3313

3301 GACTCGACACCGTGTACCTACCTGTCACCTCCCTGGGGTCACTCAGGACGCCACAGCGCA 3360  
3314 GCTGTGCGGAAGCTCCAGAGGCGCAATGACCATCTTTAAAGCTTGCAGCTGACCCAGC 3373  
3361 GCTGAGTCGGAAGCTCCCGGGAGCAGCTGACTGCTCGCTGGAGGCGCAGCCACCCGCG 3420  
3374 CTTAAGCACAGACTTTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

ULT 5  
09-572-423B-3  
sequence 3, Application US/09572423B  
atent No. 6331399  
GENERAL INFORMATION:  
APPLICANT: Brett P. Monia  
APPLICANT: William A. Gaarde  
APPLICANT: Edward Wancewicz  
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
FILE REFERENCE: ISPH-0462  
CURRENT APPLICATION NUMBER: US/09/572.423B  
CURRENT FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 3  
LENGTH: 4015  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (56)...(3454)  
09-572-423B-3

Very Match 45.8%; Score 1500.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 563 GTACCAAAATTTGTGCCACACGATATCTGGCCCTCTGTCTCGCTAGTTACAGGCCAC 622  
DB 580 GTACCAAGCTCGCGCTGTCACACTCAGCCCGGCCCCGCCACACGCTAGTGAGCCCGAG 639  
QY 623 CCACCCGCTGGGAGGAATTTTCACTAACTTAGTTTCAACAAGATCAAGAGCAGTAG 682  
DB 640 GCCTCTGGGATGC-----GAAAGGGCTGGAACCATAG 672  
QY 683 TCGCCAGGAAGACACGAAACCCCTGGCTTCCTCATCTCGAGGTACAAAGAGGCATCTGAG 742  
DB 673 CGTCAGGAGGCGCGGGTCCCTCGCTGGCTGCGAGCCCGGGTGGAGAGCGCGGGG 732  
QY 743 TCTCACCAGTACAAGTGTGCTTCAGCTAAGAGGCCAGATGCTATCTCTCTCCGAGAGT 802  
DB 733 CAGTCCAGCCGGAAGTCTGCGCTGCCCCAAGAGGCCAGGCGTGGCTGCCCCCTGAGCC 792  
QY 803 GGAGGAGGAGACCCCAAGCAGGTGCTACCAACCCCTCAGGCAATCATGGGTGCCAAG 862  
DB 793 GGAGCGAGCGCCGTTGGGCGAGGGTCTTGGGCCCAACCCCGGCGAGGACGCGTGGACCCAG 852  
QY 863 TCCTGTCTGGTCCCGAGGTGCTACT-----GCAGAGAAAGATTTGTCTTTCTAA 913  
DB 853 TGACCGTGTGTTCTGTGTGTGCTACCTGCCAGACCGCGCGAGAGCCACTCTTTTGA 912  
QY 914 AGGAAAGGTGCTGACCTGAGTCTTC-----TGGGTGGGTGCTGTCTAAACACAGCCAG 970  
DB 913 GGGTGGCTCTCTGGCAGCGCGCACTCCCACTCCGTGGGCCCGCCAGCACACGCGGG 972  
QY 971 CTCACAT---CTCTGTCTCACCACCGCCCAAAATGCTTTTTCAGCTCAGGCCATTTAT 1027  
DB 973 CCCCCCATCCACATCGCGGCACACGCTCCCTGGGACACGCTTGTCCCGGTGTAGCG 1032  
QY 1028 TGAGACCAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGCGTCTTAAACCCCTCAAT 1087  
DB 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGGCGCCCTCTT 1089  
QY 1088 CTTACTAGCAACTCTCAGCCTTACTTGTGCTGGGCGAGGACTGGTGAGATCATCTT 1147  
DB 1090 CTTACTAGCTCTCTGAGGCCAGCCTGCTGGCGCTCGAGGCTCGTGAGACCATCTT 1149  
QY 1148 TCTGGGTCAAGGCTTAGGACATCAGGACCATCTTGCAGGACACACCGCTCTATCGCGTGC 1207  
DB 1150 TCTGGGTTCAGGCTTGTGATGCCAGGACTCCCGCAGCTTGGCCCGCTGCCACAGCG 1209  
QY 1208 ATATGCGAGATGGGCGCTGTTTCAACAGCTGTGTGTGAACCATGACAGTGTGCAATA 1267  
DB 1210 CTTATGCAAAATGCGGCCCTGTTTCTGAGCTGCTTGGGAAACCAAGGCGAGTGGCCCTA 1269  
QY 1268 TGTGAGACTCTCAGGTCATTTGACAGGTTTGAACAGCAAAACCAAGGTGACAGATGC 1327  
DB 1270 CGGGGTCTCTCTAAGACGCACTGCGCGCTGCGAGTGGGTGCGGTCAGCGAGCGGTGT 1329  
QY 1328 CT-----TGAAACACAGGCC 1342  
DB 1330 CTGTGCGCGGAGAACGCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAGACACAGACCC 1389  
QY 1343 ACCGCACTCATGATTTGCTCGCTGCACAGCTTCTCTGCGAGTATATGTTTCT 1402  
DB 1350 CGGTGGCTGTGTGAGCTGTCTCCGACAGACAGACCCCTGCGAGTGTACGGTCTCT 1449  
QY 1403 TCGGCGCTGTCTCTGCAAGGTGTGTCTGTCTGTGGGTACAGGCACAATGAGG 1462  
DB 1450 CGGGCTCTGCTGCGCGGTGTGTGGCCCGCCAGGCTCTGGGGCTCCAGGCAAAAGAAAG 1509  
QY 1463 CGCTTCTTTAGAACTTAAAGATTCATCTGTTGGGGAATACGGCAAGTACACT 1522  
DB 1510 CGCTTCTCAGGAACCAAGAGTTCTCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
QY 1523 GCAGAACTGATGTGAAGATGAAAGTAGAGGATTTGCCACTGGCTCCGAGAGCCCGGG 1582  
DB 1570 CGAGAGCTGACGTGGAAGATGAGCGTGGCGGACTGGCTTGGCTGGCGAGGCCAGG 1629

1583	GAAGGACCGTGTCCCGCTGCAGAGCACCGTTTGAGGGAGAGATCTGGGTACTGTCCT	1642
1590	GATTGGCTGTGTTCCGGCCGAGAGCACCGTCTCGCTGAGAGAGATCCTCGGCAAGTTCCT	1649
1603	GTCTGGCTGATGACACATACATGTTGGTACAGTCTGTAGGTCAATCTCTTTATCATCACAGA	1702
1610	GCACCTGGCTGATGAGTGTACGTCTCGAGCTGCTCAGGTCTCTTTCTTTATGTACGGA	1709
1703	GAGCACATTCOCAGAAGAACAGGCTCTTCTTCTACCGTAAGAGTGTGTGAGCAAGCTGCCA	1762
1750	GACCACGTTTCAAAGAAACAGGCTCTTTTTTCTACCGAAGAGTGTCTGGAGCAAGTTCGA	1809
1763	GAGCATTCGAGTCAAGGCAACAACCTCTGAGAGAGTGGGCTTACGGGAGCTGTCCAAGAGGA	1822
1810	AAGCATTCGAATCAGACAGACATTTGAAGAGGTGCAGCTTCGGAGAGCTGTCEGAAGCAGA	1869
1823	GCTCAGGCATCACCAGGACACCTGGCTAGCACAATGCCCATCTGCACAGTGCAGCTTCATCCC	1882
1870	GCTCAGGCAGCATCGGGAAGCAGGCCCGCCCTGTGCTGACGTCACAGCTCCGCTTCATCCC	1929
1883	CNAGCCCAACGGCTGCGGCCCATTTGTAACATAGTTATAGCATGGGTACAGAGCTTT	1942
1930	CAAGCCTCAGCGGCTGCGGCCGATTTGTGAACATGGACTACGCTGTGGGAGCGCAGAACGTT	1989
1943	GGCGAAGAGGAGCAGGCCCGCAGCATTTCAACCAGCGCTCTCAAGACTCTCTTTCAGCATGCT	2002
1990	CGCAGAGAAAGAGGGCCGAGGCTCTCACCTCAGGGGTGAAGGCACGTGTTACGCGTACT	2049
2003	CAACTATAGCGGCAAAAACATCTCTCATCTTATGGGCTCTCTGTATCGGGTATGAATGA	2062
2050	CAACTACAGCGGGCGCGGGCCCCCGGCTCTCTGGGCGCTCTGTGCTGGGCGCTGGACGA	2109
2063	CATCTACAGGACCTTGGCGGGCTTTGTGCTGCGGTGCTGTGGACGACAGACCCAG	2122
2110	TATCACAGGGCTTGGCGCACCTTCGTCTGCTGTGCGGGCCAGAGACCGCGCCCTGA	2169
2123	GATGTACTTTTAAAGGCAATGTGACCGGGGCTAATGATGCCATPCCCCAGGGTAAAGCT	2182
2170	GCTGTACTTTGTCAAGGTGATGTGACGGCGCGGTACGACACCAATCCCCAGGACAGGCT	2229
2183	GGTGGAGGTTGTCGCAATATGATCAGGCATCTGGAGGACAGGTACTGTATCCGCGATA	2242
2230	CAGGAGGTCACTCCGAGCATATCA---AACCGAACAAGTACTCGTGTGCTCGGT	2286
2243	TGCAGTGTCCGAGAGATAGCAAGGCCAAAGTCCACAAGTCTTTAGGAGACAGGTTCAC	2302
2287	TGCCGTGTCTCAGAAAGCCGCCATGGGCACGTTCGCAAGGCTTTCAAGAGCCACGTCCTC	2346
2303	CACCTCTCTGACCTTCAGGCATACATGGGCCAGTTCCTTAAGCATCTGAGAGATTGAGA	2362
2347	TACCTTGAACACCTCAGCGGTACATCGACACAGTTCGTGGCTCACTCTGSCAG-----GA	2400
2363	TGCCAGTGTCACTGAGGAACCTCCGTTGTCACTGAGCAGAGACATCTCATGAATGAGAGCAG	2422
2401	GACCAGCCGCTGAGGGATGGCGTGTGTCATGAGCAGAGCTCTCTCTCTGANTGAGGCCAG	2460
2423	CAGCAGCTGTTTGACTTCTTCCTGCACTTCTGTGGTCACTAGTGTGTPAAAGATTGGTGA	2482
2461	CAGTGGCCTCTTCGACGCTCTTCTCACTGCTTCATGTGCCACACACGCGGTGCGCATCAGGGG	2520
2483	CAGGTGTATACGCAAGTGCAGGGGCATCCCCCAGGGCTCCAGCCTATCCACCTGCTCTG	2542
2521	CAAGTCTTAAGTTCAGTGTGCAGGGGATCCCGAGGGCTCCATCTCTCCAGCGTGTCTG	2580
2543	CAGTCTGTGTTTCGAGACATGAGGAACAAGCTGTTGTGTGAGGTGCAGCGGGATGGGTT	2602
2581	CAGCCTGTGCTACCGGCACATGAGAAACAAGCTGTTGCGGGGATTTGCGCGGACGGGCT	2640
2603	GCTTTTAGTTTTTCTGATGACTTCTGTGTGGTGAAGCTCTGCTTGGACCAAGCAAAAC	2662
2641	GCTCTCGGTTTGTGGATGATTTCTGTTGTGGTGAACACTCACTCACCCACGCGAAAAAC	2700
2663	CTTCTCTACGACCCCTGTGTCATGGCGTCTCTCAGATATGGGTGCATGATAACTTTCAGAA	2722

2701	Db	CTTCTCAGACCCCTGGTCCGAGGTGCTCGATGCTGGTGGTGAATCTTGCGGAA	2760
2723	Qy	GACAGTGGTGAACTTCCCTGTGGAGCTGTAACCTGGGTGGTGAGCTCCATACCAGCT	2782
2761	Db	GACAGTGGTGAACTTCCCTGTAGAAAGCAGGCCCTGGGTGGCAGCGCTTTGTTCAGAT	2820
2783	Qy	GCCTGTCACTGCCTGTCTCCCTGGTGTGGCTGTGGACACTCAGACTTTGAGGT	2842
2821	Db	GCGGGCCCAACGGGCTATTCCCTGGTGGCGCTGCTGCTGGATACCCGACCCCTGGAGGT	2880
2843	Qy	GTTCGTGTACTACTCAGGTATGCCCCAGACCTCAATTAGACGAGCGCTCACCTTCCAGAG	2902
2881	Db	GCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG	2940
2903	Qy	TGTCCTTCAAGCTGGGAAGACCATGCGGAACCAAGTCTCTGCGGTCTTCGGGTGAAGTG	2962
2941	Db	CGGCTTCAAGGCTGGGAGAAATGCTGTCGAAATCTTTGGGTCTTCGGGCTGAAGTG	3000
2963	Qy	TCACGGTCTATTCTTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCATCAATATATA	3022
3001	Db	TCACAGCCTGTTTCTGGATTGTGAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA	3060
3023	Qy	CAAGATCTTCCCTCTTCAGCGCTACAGGTTTCAATGCATGTGTGATTACGCTTCCCTTTGA	3082
3061	Db	CAAGATCTCTCTGCTGAGGCGTACAGGTTTCAAGCATGTGTGCTCAGCTCCCATTTCA	3120
3083	Qy	CCAGCGTGTTAGGAAGAACTTCACATTTCTTTGGGCATCATCTCCAGGCCAAGCATCTTG	3142
3121	Db	TCAGCAAGTTTGGGAAGAACCCACATTTTCTCTGCGCGTCATCTCTGACACGGCCCTCCCT	3180
3143	Qy	CTGCTATGCTATCCTTGAAGGTCAAGATCCAGGAATGACACTAAG-----GCCTC	3193
3181	Db	CTGCTACTCATCCTGAAAGCCAAAGAACGAGGATGTCGCTGGGGGCCAAGGGGCCCG	3240
3194	Qy	TGGCTCCTTTCTCTTGAAGCCGACATTTGGGTCTGCTACAGGCGCTTCTGTCTCAAGCT	3253
3241	Db	CGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTTCTGTCAAGCT	3300
3254	Qy	GGCTGCTCATTTGTCATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT	3313
3301	Db	GACTTCGACACCGGTGCACCTAGTGCACCTCTCTGGGGTCACTTCAGGACAGCCGACGCA	3360
3314	Qy	GCTGTGCGGAAGCTCCACAGAGCGCAATGATCCATTAAAGCTTGCAAGCTGACCCAGC	3373
3361	Db	GCTAGTGGGAAGCTCCCGGGACGACGCTGACTGCCCTGGAGGCCGACGCCAACCCCGC	3420
3374	Qy	CCTAAGCACAGACTTTACAGACCATTTTGGACTAA	3407
3421	Db	ACTGGCCTCAGACTTCAAGACCATCTCGACTGA	3454

## RESULT 6

```

US-09-128-354-1
Sequence 1, Application US/09128354
Patent No. 6337200
GENERAL INFORMATION:
APPLICANT: Morin, Gregg B.
APPLICANT: Garon Corporation
TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants
FILE REFERENCE: 0153989-003310US
CURRENT APPLICATION NUMBER: US/09/128,354
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 08/851,843
EARLIER FILING DATE: 1997-05-06
EARLIER APPLICATION NUMBER: US 08/854,050
EARLIER FILING DATE: 1997-05-09
EARLIER APPLICATION NUMBER: US 08/911,312
EARLIER FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: US 08/912,951
EARLIER FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: US 08/915,503
EARLIER FILING DATE: 1997-08-14

```



EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: US 08/974,549  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 08/974,584  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 09/052,864  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
EO ID NO 1  
LENGTH: 4015  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (56)...(3454)  
OTHER INFORMATION: human telomerase reverse transcriptase (hTRT) cDNA  
09-128-354-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGAGCAATGACCGCGCTCCCTGCTGTCGCGCGCGTGGCGCTCTCTCTGCTGG 82  
|||||  
40 CCGGCGCAACCCCGGATGCGCGGCTCCCGCTGCGAGCGTGGCGCTCCCTCTGCTGG 99  
|||||  
83 CAGCCGATACCGGGAGGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 142  
|||||  
100 CAGCCACTACCGGAGGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 159  
|||||  
143 CAGGCGCTTGCAACCGCGGACCGGAGATCTACCGCACTTTGGTTGCCCAATGCCCT 202  
|||||  
160 CTGGCGCTTGTTGAGCGCGGGGACCGCGGCTTTGCGGCGCTGGTGGCCCAATGCCCT 219  
|||||  
203 AGTGTGCATGCACTGGGCTCAGAGCTTCACTGCGCAACCTTTCTTTCACAGGTGTC 262  
|||||  
220 GGTGTGGTGGCTGGGACGACAGCGCGCGCGCGCGCGCGCTCTCTTCCGCGAGGTGTC 279  
|||||  
263 ATCCCTGAAGAGCTGGGCGGCTGGCGAGTCTGCGAGCGCAACAGAGAAA 322  
|||||  
280 CTGGCTGAAGAGCTGGTGGCGGAGTCTGCGAGGCTGTGCGAGCGCGCGCGAGAA 339  
|||||  
323 CGTGTGGCTTTTGGCTTTGAGCTCTTAACGAGCGCCAGAGCGCGGCTCTCCATGGCTT 382  
|||||  
340 CGTGTGGCTTTTGGCTTTGAGCTCTTAACGAGCGCGCGCGCGCGCGCTCTCCATGGCTT 399  
|||||  
383 CACTAGTAGCGTGTAGTACTTCCCAACACTGTATTGAGACCTGCGTGTAGTGG 442  
|||||  
400 CACCAACAGCTGCGCAGCTACTCTCCCAACAGGTGACCACTGCGGGGGAGCGG 459  
|||||  
443 TGCATGGATGCTACTGTTGAGCGGAGTGGGCGACGACCTGCTGCTACTGTTGGCACA 502  
|||||  
460 GCGGTGGGCGCTGCTGCTGCGCGCGTGGGCGACGACGCTGCTTCACTGCTGGCAG 519  
|||||  
503 CTGTGCTCTTTATCTTCTGGTGCCTCCCAAGCTGTGCTTACAGAGTGTGGTCTCCCT 562  
|||||  
520 CTGCGGCTCTTTGCTGGTGGCTCCCAAGCTGTGCTTACAGAGTGTGGGCGCGCT 579  
|||||  
563 GTACCAAAATTTGTGCCACCAAGATATGCGCTCTGTGCTGCTGCTTACAGGCCAC 622  
|||||  
580 GTACCAAGCTGCGGCTGCGCACTCAGCGCGCGCGCGCGCAACGCTAGTGACCCGAG 639  
|||||  
623 CCACCCGCTGGGCGAGAAATTTCACTAACTTAGTCTTTTCAACAGATCAAGAGCAGTAG 682  
|||||  
640 GCGTCTGGGATGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 672  
|||||  
683 TCGCCAGGAGACCGAAACCCCTGCTGCACTCTGAGTACAGAGGCACTCTGAG 742  
|||||  
673 CGTCAGGAGGCGCGGGGTCTCCCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGGGG 732

QY 743 TCTCAGAGTACAGGTGTGCGCTTACAGTAAGAGCCAGATGCTATCTCTGCTCCGAGAGT 802  
DB 733 CAGTGCCAGCGGAGGTCTCCCGTTGCCCAAGAGGCCAGGCGTGGCGCTGCGCTTGAGCC 792  
QY 803 GGAGGAGGACCCCAAGAGGAGTGTCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
DB 793 GGAGCGGACGCGCGTTGGGCGAGGGTCTCTGGGCCCAACCCCGGCGAGGACGCGTGGACCGAG 852  
QY 863 TCTGTCTCGGTCTCCCGAGGTCCTACT-----GCAGAGAAAGATTTTGTCTTTCAA 913  
DB 853 TGACCGTGGTTTCTGTGTGGTGTCACTGCCAGACCCGCGGAGAGCCACTCTTTTGA 912  
QY 914 AGGAAAGGTGTCTGAOCTGAGTCTCTC---TGCGTGGTGTCTCTAAACAAGCCGAG 970  
DB 913 GGTGTGGCTCTCTGGCACGCGCACTCCCAACCATCTCGTGGCGCCAGCACCAAGCGGG 972  
QY 971 CTCCAGAT---CTCTGCTGTCAACCCCGCCAAATGCTTTTACGCTCAGGCCATTAT 1027  
DB 973 CCCCCATCCACATCGCGGCCACCGTCCCTGGGACAGCGCTTGTCCCGCGGTATAGC 1032  
QY 1028 TGAGACCAGACATTTCTTTTACTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT 1087  
DB 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGCGCA---CAAGGAGCAGCTGCGGCCCTCCT 1089  
QY 1088 CTTACTCAGCAACCTCCAGCCTAACTTGACTGGGCGCAGGAGACTGTGGAGATCATCTT 1147  
DB 1090 CTTACTCAGCTCTCTGAGGCCAGCGCTGCTGGCGCTGGAGGCTCGTGGAGACCATCTT 1149  
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTGTGAGGACACACCGCTCTATCGCGTCG 1207  
DB 1150 TCTGGGTTCCAGCGCTGTGATGCCAGGACTCCCGCAGTTGCGCCGCTGCCCGAGCG 1209  
QY 1208 ATACTGCAATGGGCGCTCTTTCCTCAACAGCTGTGTGTAACCATGCAAGAGTGCATA 1267  
DB 1210 CTACTGCAAAATCGGCGCTGTTTCTGAGCTGCTTGGGAAACCAAGCGCAGTGGCCCTTA 1269  
QY 1268 TGTGAGACTCTCAGGTCACTTGCAGGTTTTCGACAGCAAAACCAAGAGTGCAGATGTC 1327  
DB 1270 GGGGTGCTCTCTTAAGAGCACTGCGCTGAGGTGCGGTCAACCCAGAGCGCGGTGT 1329  
QY 1328 CT-----TGAAACACAGGCC 1342  
DB 1330 CTGTGCGCGGAGAGGCCCAAGGCTCTGTGGCGCGCCCGAGGAGGAGGACACAGACCC 1389  
QY 1343 ACCCACTCATGATTTGCTCCGCTGCAGCAGCTGCTGCGAGTATATGTTTTCT 1402  
DB 1390 CCGTGGCTGTGTGAGCTGCTCCGCGACAGCAGACCGCTGCGAGGTGACGGTTCT 1449  
QY 1403 TCGGCGCTGTCTCTGCAAGGTGGTGTCTGTAGTCTCTGGGTACGAGCACAATGAGCG 1462  
DB 1450 GCGGCGCTGCTGCGCGGCTGGTGCGCCCGCAGGCTCTGGGCTCCAGGCAACAAGAACG 1509  
QY 1463 CCGTCTTTAAGAACTTAAAGATTCATCTGTTGGGAAATACGCAAGCTATCACT 1522  
DB 1510 CCGTCTCTCAGGAACCAAGAAATTCATCTCCCTGGGAAAGCATGCCAAGCTCTCGT 1569  
QY 1523 CGAGAACTGATGTGAAGATGAAAGTAGAGGATGCACTGGCTCCGAGCAGCGCCCGG 1582  
DB 1570 GCGAGGCTGAGTGTGAAGATGAGCTGCGGACTGGCTTGGCTGCGCAGGAGCCAGG 1629  
QY 1583 GAAGACCGTGTCCCGCTGCGAGCAACCGTCTGAGGAGAGGATCTGGCTACGTTCT 1642  
DB 1630 GGTGGCTGTCTCCGCGCGCAGAGCAACCGTCTGCGTGAAGAGATCTGGCCAAAGTTCT 1689  
QY 1643 GTTCTGGCTGATGAGACACATACGTGGTACAGCTGCTTAGGTCACTTTTACATCACAGA 1702  
DB 1690 GCACGTGCTGATGATGTGTACGTGTGCTGAGCTGCTCAGGCTCTTTTATGTACGGA 1749  
QY 1703 GAGCACATTCAGAGAACAGGCTCTTTCTTACCGTAAAGTGTGTGGAGCAAGTGTGA 1762  
DB 1750 GACCACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTGTGA 1809

1763 GAGCATTTGGAGTCAGGCAACACTTTCAGAGAGATGCGGCTACGGGAGCTGTCAACAAGAGGA 1822

1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGTGGGGAGGTGTTCGGAAGCAGA 1869

1823 GGTCAAGCATCACAGAGACACTGGCTACGCATGCCATCTGCGACATCTCGGCTTCATCCC 1882

1870 GGTCAAGCAGCATCGGGAACCAAGCCCGCCCTGTGAGCTTCAGATCCAGATCCGCTTCATCCC 1929

1883 CAAGCCAAACCGCTCGGCCCATTTGTGAACATGATGTTATAGCATGGGTACCAAGAGCTTT 1942

1930 CAAGCCTGACGGGCTGGGCCGATTTGTGAACATGACTACGTCGTGGGAGCCAGAACGTT 1989

1943 GGGCAGAGGAAGCAGAGCCCCAGCATTTCAACCACCGGTCTCAAGACTCTCTTCAGACATGCT 2002

1990 CCGCAGAGAAAGAGGGCCAGAGGTCTCACTTCAGGGTGAAGCAGCTGTTTCAGCGTGCCT 2049

2003 CAACTATGAGCGGACAAAAATCCTCACCTTATGGGGTCTTCTGTACTGGGTATGATGA 2062

2050 CAACTCAGAGGGCGGGCGGCCCGGCCCTCTGTGGCGCTCTGTGTGGCCTGGACGA 2109

2063 CATCTACAGGACTGGCGGGCCCTTTGTGCTGGTGTGGTGTGCTCTGAGACAGACACCCAG 2122

2110 TATCCACAGGGCCCTGGGGCACTTCTGCTGTGCTGTGGCGGCGCAGGACCGCGCGCTGA 2169

2123 GATGTACTTTGTTAAGCAGATGTGACCGGGCGCTATGATGTCATGCCATCCCCCAGGCTAAGCT 2182

2170 GCTGTACTTTGTCAAGGTGATGTGACGGCGGTACGACACCATCCCCCAGGACGGCT 2229

2183 GGTGAGGTTGTTGCCAATATGATAGGCACATCGGAGGACAGTACTGTATTCGCCAGTA 2242

2230 CACGAGGTCATCCCAAGCATCATCA- --AAACCAGAAACAGCTACTGCTGCTCGGTG 2286

2243 TGCAGTGTTCGGAGAGATAGCCAAAGCCAAAGTGCACAAAGTCTTTTAGGAGACAGAGTCAC 2302

2287 TGCCGTGTTCAGAAAGCGGCCCATGGGCACTGCCAAGGCTTCAAGAGCCACAGTTC 2346

2303 CACCCTCTCTGACTTCAGCCATACATGGGCCAGTTCTCTTAAGCATCTGCAGGATTCAGA 2362

2347 TACTTTGACAGACCTCCAGCGCTACATGGACAAGTTCGTGGCTCACTGCAG- ----GA 2400

2363 TGCCAGTGCACTGAGGAATCCGTTGTCATCGAGCAGAGCATCTCTATGANTGAGAGCAG 2422

2401 GACCAGCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAG 2460

2423 CAGCAGCCTGTTGACTTCTTCCTGCACTTCTCGCTCACTGCTGCACAGTGTCTGAAGATGTGTGA 2482

2461 CAGTGGCTCTTCGACGCTTCTCCTACGCTTCATGTGCCACACACGCCGTGGCATCAGGG 2520

2483 CAGGTGCTATACGAGTGCAGAGGCATCCCAAGGCTCCAGGCTTCAGCCTATCCACCTGCTCTG 2542

2521 CAAGTCTTACGTCAGTGCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2580

2543 CAGTCTGTGTTTCGAGACATGAGAACAAAGCTGTTGCTGAGGTGACGGGATGGGTT 2602

2581 CAGCCTGTGCTACGGCGCATGAGAACAGCTGTTTGGGGGATTCGGCGGGACGGGCT 2640

2603 GCTTTTACGTTTGTGATGACTTTCTGTTGTTGACGCTCTCACTTGGACCAAGCAAAAC 2662

2641 GCTCCTCGGTTTGGTGGATGATTTCTTGTGTTGTCACCTCACTCCACCCACGCGAAAC 2700

2663 CTTCCTCAGACACCTGTGTCATGCGCTCTCTGAGTATGGGTGCATGATAAACTTGCAGAA 2722

2701 CTTCCTCAGAACCTTGCTCGAGGTCTCCTCTGAGTATGGCTCGGTGTAATCTTTCGGAA 2760

2723 GACGTGGTGAATCTCCCTGTGGAGCCTCGTACCTCGGTGGTGTGAGCTCCATACCACT 2782

2761 GACAGTGGTGAATCTCCCTGTAGAAGACGAGGGCCCTGGGTGGCAGGCTTTTGTTCAGAT 2820

2783 GCCTGCTCACTGCTCTGTTTCCCTGGTGTGGCTGTGCTGTGACACTCAGACTTTGGAGGT 2842

2821 GCGGCCCAACGGCCTATTCCCTGTGTGCGGCTGTGCTGTGATACCCGAGCCCTCGGAGGT 2880

2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAGAAGCAGGCTCACTTCCACAG 2902

2881	DB	GCAGAGCGACTACTCCAGCTTATGCCGGGACTTCATCAGACCGAGTCTCACCTTCAACCG	2940
2903	QY	TGTTCTTCAAGAGCTGGGAAGACCATCGGAACAAGCTCCTGTTCGTTCGGTTCGGGTGAAGTG	2962
2941	DB	CGGCTTCAAGGCTGGGNGGNAATGGTCGCAAACTCTTTGGGGTCTTCGGGCTGAAGTG	3000
2963	QY	TCACGGCTTATTTCTAGACTTGGAGGTGAACAGCTCCAGACAGTCTGCATCATATATA	3022
3001	DB	TCAGAGCTGTTTCTGGATTGTCAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA	3060
3023	QY	CAAGATCTTCCTGCTTCAGGCTTCACAGGTTCCATGCATGTGTGATTCAGTTCCTCTTGA	3082
3061	DB	CAAGATCTCTCTGCTGAGGGCTACAGGTTTCAGCAATGTGTGTGAGCTCCCATTTCA	3120
3083	QY	CCAGCGTGTTAGGAAGACTCATATCTTTCTGGGCATCATCTCCAGGCCAAGCATCCTG	3142
3121	DB	TCAGCAAGTTTGGGAAGAACCCACATTTTCTCGCGCGTCACTCTGACACGGCTCCCT	3180
3143	QY	CTGTATGTCTATCCTGAAGTCAAGATCCAGGAATGACATTAAG-----GGCTC	3193
3181	DB	CTGTACTCCATCTGAAGCCCAAGAACGACGGGATGCGCTGGGGGCCAAGGGCGCGCGC	3240
3194	QY	TGGTCTCTTTCTCTGAAGCCGCAATTTGGCTCTGTCTACCAAGCGCTTCTGTGCTCAAGCT	3253
3241	DB	CGGCCCTCTGCCCTCCGAGGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT	3300
3254	QY	GGCTGCTCATTTCTGTGATCTTACAAATGTCTCTGGGAAGCTCTGAGGACAGGCCCAAAACT	3313
3301	DB	GACTCGACACGGTGTCACTTACGTGCGCACTCTCTGGGGTCACTCTAGGACAGGCCACAGCGCA	3360
3314	QY	GCTGTGCGGAAGCTCCAGAGCGCAATGACCATCTCTTAAAGCTGCAGCTGACCCAGC	3373
3361	DB	GCTGAGTCGGAGACTCCCGGGGACGAGCTGACTGCCCTGGAGGCGCGCAGCAACCGCGC	3420
3374	QY	CCTAAGCACAGACTTTCAGACCATTTTGGACTAA	3407
3421	DB	ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA	3454

## RESULT 7

```

RESOL 7
US-09-675-321-1
; Sequence 1, Application US/09675321
; Patent No. 6440735
; GENERAL INFORMATION:
; APPLICANT: Gaeta, Federico C.A.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune
; Response to a Telomerase Antigen
; FILE REFERENCE: 015389-003500FC
; CURRENT APPLICATION NUMBER: US/09/675,321
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/112,006
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: WO 92/069898
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(3454)
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)
US-09-675-321-1

```

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 58.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCGCGCTTGAGCACAATGACCGCGCTCTCGTTGCCCGCGTGCCTCTGCTGCG 82



2183 GGTGAGGTTGTTGCCATATGATCAGGCACTCGAGAGCAGCTACTGTATCGCCAGTA 2242  
2230 CACGAGGTCATCGCCAGCATCA---AACCCAGAAACAGCTACTGCGTGGTGA 2286  
2243 TGCAGTGTCCGGAGAGATAGCCAAAGGCCAAGTCCCTTTAGGAGACAGGTCAC 2302  
2287 TGCCTGTGTCAGAGGCGCCCATGGGACAGTCCGCAAGGCTTCAAGAGCCACGTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAGCATCTGAGATTCAGA 2362  
2347 TACCTTGAAGACCTCCAGCCGATACATGCGACAGTTCGTGGCTCACTGCGAG-----GA 2400  
2363 TGCCAGTGCATCGAGAACTCCGTTGTTCATCGACGAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGTGGGGATGCGTGCATCGACGAGAGCTCTCTCCCTGAATGAGCCAG 2460  
2423 CAGCAGCTGTTTGAATCTCTCTGCACTTCTGCTGTCAGAGTGTGTAAGATTGGTGA 2482  
2461 CAGTGGCTCTTCGACGCTTCTCTACGCTTCTATGTCACACGCGGTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGAGTCCAGGCGATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAGTCTACGTCAGTCCAGGGGATCCCGAGGCTTCTCTCTCAGCTGCTCTG 2580  
2543 CAGTCTGTTTTCGAGACATGAGAAACAAGCTGTTTCTGAGGTGCGCGATCGGTT 2602  
2581 CAGCTGTGCTACGCGACATGAGAAACAAGCTGTTTTCGCGGATTCGCGGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGTTGTCAGCGCTCACTTGGACCAACAAC 2662  
2641 GCTCTGCTTGTGAGATGATTTCTTGTGTCACACCTCACTACCCACCGCAAC 2700  
2663 CTCTCTAGACCTCTGTCATGCGCTTCTGAGTATGGTGTGATGATMACTTCAGAA 2722  
2701 CTCTCTAGGACCTCTGTCAGGCTGCTCTGATGATGGTGTGATGATGCGAA 2760  
2723 GACAGTGTGATCTCTCTGAGCTGATGATGATGATGATGATGATGATGATGATGAT 2782  
2761 GACAGTGTGATCTCTCTGAGCTGATGATGATGATGATGATGATGATGATGATGAT 2820  
2783 GCTCTCTCACTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2842  
2821 GCGGCGCCAGGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2880  
2843 GTTCTGTGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2902  
2881 GCAGAGCACTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
2903 TGTCTTCAAGCTGGAAGACATGCGGAAACAGCTCTCTGCTCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGGAAGACATGCGGAAACAGCTCTCTGCTCTGCTGCTGCTGCTGCT 3000  
2963 TCACGCTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3022  
3001 TCACAGCTGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
3023 CAAGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3082  
3061 CAAGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120  
3083 CCAGCTGTGAGGAAGACCTCAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
3143 CTGCTATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3193  
3181 CTGCTACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
3194 TGGCTCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3253  
3241 CGGCTCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300

QY 3254 GGCTGCTCATCTGTCTATCAAAATGTCTCTGGACCTCTGAGGACAGCCAAAATCT 3313  
DB 3301 GACTCGACACCGTGTACCTTACGTGGTCCACTCTCTGGGTCACTCAGGACAGCCAGCA 3360  
QY 3314 GCTGTGCGGAAGCTCCAGAGGCGCAATGACCAATCTTTAAAGCTGCAAGTGCACCCAGC 3373  
DB 3361 GCTGAGTCCGAAGCTCCGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
QY 3374 CCTAAGCACAGACTTTTCAGACCACTTTTGGACTAA 3407  
DB 3421 ACTGCTCAGACTTCAAGACCATCTCTGACTGA 3454

## RESULT 8

US-09-052-919-1  
; Sequence 1, Application US/09052919  
; Patent No. 644650  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Antisense Compositions for Detecting and  
; TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/052,919  
; FILING DATE: 31-MAR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/974,549  
; FILING DATE: 19-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/974,584

FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "human telomerase reverse transcriptase (hTRT)"  
-09-052-919-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 23/9; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGGCTTTGAGCACAATGACCGCGCTCTCGTTGCGCGCGCGTGGCGCTCTCTGCTGGC 82  
40 CCGGGCCACCCCGCGATCGCGGCGCTCCCGCTGCCGAGCGTGGCTCTCTGCTGGC 99  
83 CAGCGGATACCGGAGGTGGCGGCTGGCAACCTTTGTGGCGCGCTGGGCGCGGAGG 142  
100 CAGCCCACTACCGGAGGTGGCGGCTGGCAACCTTTGTGGCGCGCTGGGCGCGGAGG 159  
143 CAGCGGCTTTGCAACCGCGGACCGCAAGATCTACCGCACTTTGGTTGCCAATGCT 202  
160 CTGGCGCTGGTACGCGGGGACCCGCGGCTTTCCGCGCGCTGGTGGCGCGCT 219  
203 AGTGTGATGCACTGGGGTCAAGCTTCACTGCGCACTTTCCGCACTTTCCACAGGTC 262  
220 GGTGTGCTGCTGGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279  
263 ATCCCTGAAGAGCTGGTGGCGAGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCTGAAGAGCTGGTGGCGCGAGTGTGAGAGGCTGTGAGAGCGCGCGCGAGAA 339  
323 CGTGTGCTTTTGGCTTTGAGTGTAAAGAGTCAAGAGCGCGCGCGCTCCCATGGCTT 382  
340 CGTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 399  
383 CACTAGTAGCTGCTAGTACTTGGCGGCACTGTATTGAGACCTGCTGCTGCTGCTGCT 442  
400 CACACAGCGTGGCGAGTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 459  
443 TGCATGATGCTACTGTGAGCGAGTGGCGGCGAGCACTGCTGCTGCTGCTGCTGCTGCT 502  
460 GCGTGGGGCTGCTGCTGCGCGCGTGGCGGCGAGCGTGTGCTGCTGCTGCTGCTGCT 519  
503 CTGTGCTTTTATCTTGTGTCGCCCGCGAGTGTGCTTACGAGTGTGGTGTCTCCCT 562  
520 CTGCGCGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAATTTGGCCACGAGTATCTGGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639  
623 CCGACCGTGGCGAGGAATTTCTAATCTTAGTGTCTTACCAAGATCAAGAGCAGTAG 682

Db 640 CGGTCTGGATGC-----GAAAGGCTTGAACATAG 672  
QY 683 TCGCAGGAAGCACCAGAACCCCTGGCTTGCATCTGAGGTACAAAGAGCATCTGAG 742  
Db 673 CGTCAGGAGGCGCGGGTCCCTTGGCTTGCAGGCGCGGGTGCAGGAGCGCGGGG 732  
QY 743 TCTCAGCAGTACAAAGTGTGCTTACGCTAAGAGGCGAGATGCTATCTGTCTCCAGAGT 802  
Db 733 CAGTGCCAGCGGAAGTCTGCGTGGCCCAAGAGCGCCAGGCTGGCGCTGCCCTGAGCC 792  
QY 803 GGAGAGGAGCACCACAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGGTGCAAG 862  
Db 793 GGAGCGAGCGCGCTTGGCGAGGGTCTTGGCGCCACCCCGGCGAGGACGCGTGCACGAG 852  
QY 863 TCCTGCTCGTCCCGCGAGGTGCTACT-----GCAGAGAAAGATTGTCTCTTAA 913  
Db 853 TGACGCTGTTTCTGTGTGTGTGTCCTGTCAGACCGCGCGAAGAGCCACCTCTTTGA 912  
QY 914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGGTGTGTGTAAACACAAGCCAG 970  
Db 913 GGTGGCGTCTCTGCGACGCGCCACTCCACCCATCGTGGCGCGCAGCACCCACGCGG 972  
QY 971 CTCACAT---CTGTGTGTACCAACCCCGCGCAAAATGCTTTCAGCTCAGGCCATTAT 1027  
Db 973 CCCCCCATCCCATCGCGCGCCACCATCTCCCTGGGACACGCTTGTCCCGCGGTGTACGC 1032  
QY 1028 TGAGACCAAGCATTTCTTTACTCCAGGGAGATGGCCAGAGCGTCTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCATTTCTTACTCTCAGGCGA---CAAGGACAGCTGCGGCGCTCTT 1089  
QY 1088 CTTACTCAGCAACTCCAGCTTAACTTGTGCTGGGCGAGAGACTGGTGGAGATCATCTT 1147  
Db 1090 CTTACTCAGCTCTCTGAGGCGCGAGCTGCTGAGGCTCTGAGAGCTCTGAGAGCATCTT 1149  
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGACCACTCTGCGAGACACCGCTCTATCGGCTG 1207  
Db 1150 TCTGGTTTCCAGGCGCTTGTATGCGAGGATCTCCCGAGGTTGCCCGCTTGCCTCCAGCG 1209  
QY 1208 ATACTGGCAGATCGGCGCTTGTTCACAGCTGTGGTGAACCATGCGAGGTGCCAATA 1267  
Db 1210 CTACTGGCAATCGGCGCTTGTTCGAGGCTGTCTGGGAGACCGCGCGAGTGCCTCTA 1269  
QY 1268 TGTGAGACTCTCAGGTCACATTCAGGTTTTCGAGACAGCAACCAACAGGTGACAGATGC 1327  
Db 1270 CGGGGTGCTCTCTCAAGAGCGACTGCGGCTGCGAGTTCGCGTCAACCCAGCGCGGTGT 1329  
QY 1328 CT-----TGAAACACAGCCC 1342  
Db 1330 CTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGGCGCGCGAGGAGGAGCACAGACCC 1389  
QY 1343 ACCGACCTCATGATTTGCTCCGCTGCGACAGAGTCCCTGGCGAGTATATGTTTCT 1402  
Db 1390 CGCTGCTGTGTGAGTGTCTCCCGCAGCAGAGCGCTTGGCGAGTGTACGGCTTCTGT 1449  
QY 1403 TCGGCGCTGTCTCTGAGAGTGTGTCTGTAGTCTCTGGGGTACCAGGCACAATGAGCG 1462  
Db 1450 GGGGCGCTTGCCTGCGCGGCTGTGTGCCCGAGGCTCTGGGCTCCAGGCAACAGAACG 1509  
QY 1463 CGCTTCTTTAAGAACTTAAAGAGTTCATCTCGTGGGAAATACGGCAAGTATCACT 1522  
Db 1510 CGCTTCTCTCAGGAACACCAAGAGTTTCTCTCCCTGGGAAAGCATGCCAAGCTCTGCT 1569  
QY 1523 GCAGAACTGATGTGAAGATGAAAGTAGAGATTGCCACTGGCTCCGCGAGAGCGCGGG 1582  
Db 1570 GCGAGGCTGAGTGAAGATGAGGTCGCGGACTGGCTTGGCTGCGCAGGAGCGCGAG 1629  
QY 1583 GAAGACCGTGTCCCGCTGCGAGCAACCGTCTCAGGAGAGGATCTCGGTACGTTCTCT 1642  
Db 1630 GGTGGCTGTGTTCGCGCGCAGAGCAACCGTCTCGTGAAGAGATCTCGGCAAGTTCTCT 1689  
QY 1643 GTTCTGGCTGTGAGACATACGTTGATAGCTCTTAGGTCTTCTTTTACATCAGAG 1702

1690 GCACCTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGTCTCTTTCTTTATGTACAGGA 1749  
1703 GAGCACATTCAGAGAACAGAGCTCTTTCTTCTACCGTAAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAGAGCTCTTTCTTCTACCGTAAGAGTGTGTGAGCAAGTTCGA 1809  
1763 GAGCATTTGGAGTTCAGGCAACACCTTTGAGAGGTGCGGTACGGAGCTGTCAAGAGGA 1822  
1810 AAGCANTTGAATCAGACAGCACTTGAAGAGGTGAGCTCGGAGAGCTGTGAGAGCAGA 1869  
1823 GGTGAGGCATCACCAGGACACCTTGGCTAGCCATGCCCCATCTGACAGACTGCGTTTCATCCC 1882  
1870 GGTGAGGCAGCATCGGGAAGCAGGCGCGCTGCTGACGTCCAGACTCGCTTCATCCC 1929  
1893 CAAGCCCAACGGCTGCGGCCCATTTGTGAACATGAGTTATGACATGAGTACGAGACTTT 1942  
1930 CAAGCCTGACGGGTGCGGCCGATTTGTGAACATGAGTACGCTGCGGAGCCAGAACGTT 1989  
1943 GGCAGAGAGGAGCAGGCGCCAGCAATTTCAACCAGCGTCTCAAGACTCTCTTCAGCAATGCT 2002  
1990 CGCAGAGAAAGAGGGCCGAGCGTCTCACTCGAGGGTGAAGGCACTGTTCAAGCGTGTCT 2049  
2003 CAACTATGAGCGGCAAAACATCTCACTTATGAGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACTACAGCGGGCGCGCGCCCGCCCTCTGCGGCGCTCTGTGTGCGGCGCTGAGCA 2109  
2063 CATCTACAGGACCTGGCGGGCTTTGTGCTGCTGCTGCTGCTGAGCAGACACCCAG 2122  
2110 TATCCACAGGGCTGGCGCACTTCTGTGCTGCTGCTGCTGCGGCCAGGACCGCGCTGA 2169  
2123 GATGATCTTTGTAAGCAGATGTGACCGGGGCTTATGATGCCATCCCGAGGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGATGTGACGGGCGCTACGACACCATCCCGCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCGAGTA 2242  
2230 CAGGAGGTCTATCGCCAGCATATCA--AACCACAGAACAGTACTGTGCTGCTGCTGTA 2286  
2243 TGCAGTGTGCGGAGATAGCCAGGCGAAGTCCACAAGTCTCTTTAGGAGACAGGTCTAC 2302  
2287 TGCCTGTGTCCAGAGGCGGCCATGGGACGTCCGCAAGGCTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTCTTAAAGCATCTGAGAAATTCAGA 2362  
2347 TACCTGTGACAGACCTCCAGCGGTACATGCGAGTTCGTGGCTCACTGCGAG-----GA 2400  
2363 TGCCAGTGCATGAGGAACCTCGTGTGTCATCGAGCAGAGCATCTCTATGAATCAGAGCAG 2422  
2401 GACCAGCCGCTGAGGGATGCGTGTGTCATCGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGTGCTCTCTGCACTTCTGCTGCAAGTGTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTCGACGCTCTCTCTACGTTCTCATGTCACACCGCTGCGCATCAGGGG 2520  
2483 CAGGTGCTATACGAGTGCAGGGCATCCCCAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAACTCTACGTCAGTGCAGGGATCCCCAGGCTCTCATCTCTCTCTCTCTCTCTCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAACAGCTGTTGCTGAGGTGAGCGGATGGGTT 2602  
2581 CAGCCTGTGCTACGCGCAGATGAGAACAGCTGTTGCGGGATTCGGCGGACCGGCT 2640  
2603 GCTTTTACGTTTGTGATGATTTCTGTTGTCAGCGCTCTCTCTCTCTCTCTCTCTCTCT 2662  
2641 GCTCTGCTGTTGTTGATGATTTCTGTTGTCAGCTCTCTCTCTCTCTCTCTCTCTCTCT 2700  
2663 CTTCTCTCAGCACCCTGGTCCATGGGTTCTCTGATGAGGTGCTGATGAATAAATTCAGAA 2722  
2701 CTTCTCTCAGCACCCTGGTCCAGGTGTCCTCTGATGATGCTGCTGCTGCTGCTGCTG 2760  
2723 GACAGTGTGAACTTCTCTCTGAGCCCTGTGACCTGGTGGTGGTGCAGCTCATACAGCT 2782  
2761 GACAGTGTGAACTTCTCTCTGATGAGACAGGCGCTCTGTTGCGACGCTTTGTTTCAGAT 2820

Qy 2783 GCCTGCTCACTGCTCTTTCTCTGCTGTGGCTTGTGCTGAGCACTCAGACTTTTGAGGT 2842  
Db 2821 GCGGGCCCAAGGCTTATCTCTGCTGTGGCTTGTGCTGAGTATCCCGACCTTGAGGT 2880  
Qy 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACCTTCCAGAG 2902  
Db 2881 GCAGAGCACTACTCCAGCTATGCCGAGCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
Qy 2903 TGTCTTCAAAGCTGGGAAGACCATGCGGAAACAGCTCCTGTGCTGCTTGTGGTGAAGTG 2962  
Db 2941 CGGCTTCAAGGCTGGGAGAACATGCTGCGCAACTCTTTTGGGGTCTTGTGGCTGAAGTG 3000  
Qy 2963 TCAGGCTTATTTCTAGACTTGCAGGTGAACAGCTTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGATTTGAGGTGAACAGCTTCCAGACGCTGTCACCAACATCTA 3060  
Qy 3023 CAAATCTTCTCTGCTCAGGCTTACAGGTTCAAGTTCATGATGTGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAATCTTCTCTGCTCAGGCTTACAGGTTTCAAGTTCATGATGTGTGCTGCTGCTCCTTCA 3120  
Qy 3083 CCAGCTGTTAGGAAGAACTCACAATTTCTTCTGGGCTATCTCCAGCAAGCATCTG 3142  
Db 3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTGCGCTCATCTCTGACGCTCCTCT 3180  
Qy 3143 CTGCTATGCTATCTGAGGTCAAGATCCAGAAATGACACTAAAG-----GCCTC 3193  
Db 3181 CTGCTACTCTCTGAAAGCCAAAGACGAGGATGCTGCTGGGGGCCAAGGGCGCGC 3240  
Qy 3194 TGGCTCTCTTCTCTGAGGCGCACATTTGGCTTGTCTACGAGGCTTCTCTCTCAAGCT 3253  
Db 3241 CGGCTCTCTCTCTGAGGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
Qy 3254 GGTGCTCATCTCTGCTCATCTACAAATGTCTCTGGGAGCTCTGAGGACAGCCCAAACT 3313  
Db 3301 GACTCGACACCGTGTCACTACGTGCCACTCTCTGGGTCACTCAGGACAGCCAGACGCA 3360  
Qy 3314 GCTGTGCGGAAGTCCAGAGCGCAATGACCATCTTAAAGTGTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTGGAACTCCCGGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
Qy 3374 CCTAAGCACAGACTTTTCAAGCACTTTTGGACTAA 3407  
Db 3421 ACTGCTCTGAGCTTCAAGCACTTCTGAGCTGA 3454

## RESULT 9

US-08-912-951-1  
; Sequence 1, Application US/08912951  
; Patent No. 6475789  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
; TITLE OF INVENTION: THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 335  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
OTHER INFORMATION: /note= "human telomerase reverse  
OTHER INFORMATION: transcriptase (hTERT) catalytic protein  
OTHER INFORMATION: component"

-08-912-951-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGGGCTTTGAGCAAAATGACCCGGCTTCTGCTGGTGGCCCGGTGGCTCTCTGTGG 82  
40 CCGGGCAACCCCGGAGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 99  
83 CAGCCGATACCGGAGGTGTGCGCTGCGCACTTTGTGGGGCGCTGGGGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 159  
143 CAGGGCGCTTGTCAACCCGGGAGCCGAGATCTACCGCACTTTGTGTGCGCAATGCT 202  
160 CTGGCGCTGTGTGAGCGGGGACCCCGCGCTTTCGCGCGCTGTGTGCGCGCTGTG 219  
203 AGTGTGATGCTGGGGTACAGCTGCGACCTGCGACCTTTCTTCCACCGAGGTGTC 262  
220 GGTGTGCTGGTGGGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
263 ATCCCTGAAGAGCTGGTGGCCAGGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCTGAAGAGCTGGTGGCCGAGGTGTGCGAGAGCTGTGCGAGCGCGCGCGAGAA 339  
323 CGTGTGGCTTTTGGCTTGTGAGCTCTTAAAGGCGCAGAGCGCGCGCTCCCATGGCTT 382

Db QY 340 CGTGTGGCTTTCGGCTTCGCGCTGCTGGAGCGGGCCCGCGGGGGCCCCCGAGGCGCTT 399  
QY 383 CACTAGTAGCGTCTAGTACTTGGCCCAACTTATTATGAGACCTGCGTGTCTAGTGG 442  
Db 400 CACCACCGAGCGTGGCAGCTACTGTCGCCAACAACGCTGACCGACGACTGCGGGGAGCGG 459  
QY 443 TGCATGGATGCTACTGTTGAGCGGAGTGGGCGACACCTGTGTGGTCTTACCTGTCTGGCACA 502  
Db 460 GGGTGGGGGCTGCTGCTGGCGCGCGGTGGCGACGACGCTGCTGGTTTCACTGCTGGCAG 519  
QY 503 CTGTGCTCTTTATCTTTCTGGTGGCCCGCAGCTGTGCTACAGAGTGTGTGGTCTTCCCT 562  
Db 520 CTGCGCGCTCTTTGTGCTGGTCTGCCAGCTGCCCTACCAAGTGTGCGGGCCCGCT 579  
QY 563 GTACCAAAATTTGTGCCACACGATATCTGGCCCTCTGTGTCCGTAGTTACAGGCCAC 622  
Db 580 GTACAGCTCGGCGCTGCCACTCAGGCCCGCGCCCGCCACACGCTAGTGGACCCGAG 639  
QY 623 CCGACCGCTGGGCGAGAAATTTCACTAACCTTAGTCTTACCAACAGATCAAGAGCAGTAG 682  
Db 640 GCGTCTGGGATGC-----GAACGGGCGCTGGAACCATAG 672  
QY 683 TCGCCAGGAGCACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
Db 673 CGTCAGGAGGCGCGGTCTCCCTGGCTTGCAGCCCGCGGTGCGAGGAGCGCGGGGG 732  
QY 743 TCTCACCAGTACAAGTGTGCTTTCAGCTAAGAGGCGAGATCTATCTCTGTCTCCGAGAGT 802  
Db 733 CAGTGCCAGCGGAGTCTGCGCTTGCCTCAAGAGGCGCGCTGCGGTGCTCCCTGAGCC 792  
QY 803 GGAGGAGGAGCCCCCAGCAGGAGTGTCTACCAACCCATCAGGCAATCATGGTGGCAG 862  
Db 793 GGAGCGGAGCCCCGTGGGCGAGGGGTCTTGGGCGCACCCCGCGCAGACGCGTGGACCGAG 852  
QY 863 TCCTGCTCGGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTCTCTCTTAA 913  
Db 853 TGACCGTGGTTTCTGTGTGGTGTGCTGCCAGACCCCGCGAAGAGCACCTTTTGA 912  
QY 914 AGGAAAGGTGTGACCTGAGTGTCTC-----TGGGTGGTGTGCTGTAAACAAGCCAG 970  
Db 913 GGGTGGCTCTCTGGCAGCGCGCACTCCCAACCCATCCGTGGGCGCGCAGCACCGCGGG 972  
QY 971 CTCACAT---CTCTGTGTACACACCCCGCCAAATGCTTTTCAGCTCAGGCAATTTAT 1027  
Db 973 CCCCCATTCACATCGCGGCGCACAGTCTCTGGGACAGCCTTGTCTCCCGGTGTACGC 1032  
QY 1028 TGAGACAGACATTTCTTTACTTCCAGGGGAGATGCCAAGAGCGCTTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCATTCTCTTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
QY 1088 CTTACTCAGCAACCTCCAGCCTAACTTGACTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147  
Db 1090 CTTACTCAGTCTCTGAGGCGCCAGCCTGACTTGGCGCTCGAGGCTCGTGGAGACCATCTT 1149  
QY 1148 TCTGGGCTCAAGGCTAGGACATCAGACCCACTCTCTCAGGACACACCGCTCTATCGGCTG 1207  
Db 1150 TCTGGTTCCAGGCGCTGGATGCCAGGACTTCCCGCAGGTTGCCCGCTTGCCTCCAGG 1209  
QY 1208 ATACTGGAGATGGGCGCCCTGTTCGAAGCTGCTGTGAACCATGCGAGATGGCAATA 1267  
Db 1210 CTACTGGCAATGCGGCGCCCTGTCTTGGAGCTGTGGAACCCAGCGCAGTGGCCCTTA 1269  
QY 1268 TGTCAAGACTCTCAGGTCAATTTGCAAGTTTCGAACAGCAAAACCAACAGGTGACAGATGC 1327  
Db 1270 CGGGGTGCTCTCTCAAGACGCACTGCTCCGCTGCGAGTGCCTTCAACCCAGGCGGTGT 1329  
QY 1328 CT-----TGAAACACGAGCC 1342  
Db 1330 CTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGGCGCCCGAGGAGGAGGACACAGACC 1389  
QY 1343 ACCGACCTCATGATTTGCTCCGCTGACAGCAGTCTCCCTGGCAGGTATATGGTTTCT 1402  
Db 1390 CCGTGGCTGTGTGAGCTGTCTCGCCAGACAGAGCCCTTGGCAGGTGTACGGCTCTGT 1449

1403	TCGGCCCTGTTCTCTGCAAGGTGGTGTCTGTAGTCTCTGGGTACAGGCACAATGAGCG	1461
1450	GGGGCCCTGCTGGCGCGGTGGTGGCCCCAGGCCCTCTGGGCTCAGGCACAACGAAG	1509
1463	CGCTTCTTTTAAAGACTTAAAGAGTTTCATCTGTTGGGGAATACGCCAGCTATCACT	1522
1510	CGCTTCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT	1569
1523	GCAAGAACTGATGTGGAGAGTAAAGTAGAGGATGCCACTGGCTCCGACAGCCCGGG	1582
1570	GCAAGAGCTGACGTGGAGATGAGCGTGGGGACTCGCTCTGGCTGCGCAGAGGCCAGG	1629
1583	GAAGACCGTGTCCCGCTGCAGAGACCGCTCTGAGGAGAGAGTCTGGCTAGTTCCT	1642
1630	GTTGGCTGTGTTCCGGCCGAGAGACCGCTCTGGCTGAGAGAGTCTGGCCAAAGTTCTT	1689
1643	GTTCTGGCTGATGGACACATACGTGGTACAGCTGCTTAGGTCAATCTTTTATCATCACAGA	1702
1690	GCACTGGCTGATGAGTGTGTAAGCTCGTGCAGCTGCTCAGTCTCTTTTATGTCAAGGA	1749
1703	GAGCACATTCAGAAAGAACAGGCTCTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA	1762
1750	GACCACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGACAAAGTTGCA	1809
1763	GAGCAATCGAGTCAAGGCAACACCTCTGAGAGAGTGGGCTTACGGAGAGTGTCCACAGAGGA	1822
1810	AAGCATTTGGATTCAGACAGCACTTGAAGGGTGCAGTCTCGGAGAGTGTTCGGAAGAGA	1869
1823	GGTCAGGCATCACAGAGACACTTGGCTAGCCATGCCCATCTGCAGACTGCGCTTCATCCC	1882
1870	GGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCC	1929
1883	CAAGCCCAACGGCTCGGGCCCATTTGACATGAGTATACATGGGTACAGAGCTTT	1942
1930	CAAGCCTGACGGCTCGGGCCGATTTGGAACATGGACTAGTCTGTGGGAGCCAGAAAGTT	1989
1943	GGCGAAGAGGAGCAGGCCAGCATTTTCAACCAGCGCTCTCAAGACTCTTCTTCAGCATGCT	2002
1990	CCGACAGAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACATGTTTCAGCGTCT	2049
2003	CAACTATGAGGGGCAAAACATCCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA	2062
2050	CAACTACAGCGGGCGCGGGCCCGCCCTCTCTGGCGCTCTGTGCTGGGCGCTGGACGA	2109
2063	CATCTACAGGACCTTGGCGGGCTTTGTGCTGGCTGTGCTGTCTGGACAGACACCCAG	2122
2110	TATCCACAGGGCTTGGCGCACTTCTGTGCTGGCTGTGGGGCCAGAGACCCGCGCTGA	2169
2123	GATGTACTTTTAAAGCAATGTGACCGGGCTTATGATGCCATGCCATCCCCAGGTAAGCT	2182
2170	GCTGTACTTTGTCAAGGTGATGTGACGGCGCGTACGACACATCCCCCAGGACAGGCT	2229
2183	GGTGGAGTGTGTGCCAATATGATCAGGCATCTCGGAGACAGTACTGTATCCGCCAGTA	2242
2230	CACGAGGTCTATCCCAAGCATATCA----AACCCAGAAACAGTACTCGTGGTGGTCCGTA	2286
2243	TGCAGTGTCCGGAGAGATAGCCAAAGCCAAAGTCCACAAGTCTTTTAGGAGACAGGTCAAC	2302
2287	TGCCTGTGTCAGAAAGCCGCCCATAGGCACGTCGCAAGGCCCTTCAAGAGCCACGTCCTC	2346
2303	CACCTCTCTGACCTTCAGGCATACATGGGCCAGTTCCTTAAAGCATCTGCAGGATTCAGA	2362
2347	TACCTTTGACAGACCTTCAGCCGTACATGGACAGTTCGTGGCTCACTCTGAG-----GA	2400
2363	TGCCAGTGCATGAGGAACCTCGTGTGTCATCCAGACAGACATCTCTATGAATGAGACGAG	2422
2401	GACCAGCCGCTGAGGGATGCCGTGCTCATCGAGCAGAGCTCTCTCCCTGATATGAGGCCAG	2460
2423	CAGAGACCTGTTGACTTCTTCCTGCACATCTCTGCGTGCACAGTGTGTAAGAATTTGTGTA	2482
2461	CAGTGGCCCTCTTCGACGCTTTCTCTCATGTGCCATCATGTGCCACCAACGCGCTGGCATCAGGG	2520

Qy	2483	CAGGTGCTATACGAGTGCACGAGCATCCCCACGAGGCTCCAGAGCTATCCACCTGCTCTG	2542
Db	2521	CAAGTCCCTACGTCAGTGCACGAGGATCCCGCAGAGGCTCCATCTCTCCACGCTGCTCTG	2580
Qy	2543	CAGTCTGTGTTTCCGAGACATGAGAGAACAGCTGTTTGTCTGAGGTGACGCGGATGGGTT	2602
Db	2581	CAGCCTGTGCTACGCGACATGAGAGAACAGCTGTTTTCGGGGATTTCCGCGGACGGGCT	2640
Qy	2603	GCCTTTACGTTTGTGTGATGACTTTCTGTGTGTGACGCTCATCTTGGACCAAGCAAAAC	2662
Db	2641	GCTCTCGGTTTGTGTGATGATTTCTTTGTGTGTGACCTCATCTCACCTCACCCACGCGAAAC	2700
Qy	2663	CTTCTCTCAGCACCCCTGTCTCATGGCGTCTCCTCAGTATGGGTGCATATAAACTTGCAGAA	2722
Db	2701	CTTCTCTCAGGACCTCTGTCCGAGGTGTCCCTGAGTATGGCTGGTGGTGAATTTGCGAA	2760
Qy	2723	GACAGTGTGAACTTCCCTGTGAGGCTCGGTAACCTGGGTGGTGCAGCTCCATACCAAGCT	2782
Db	2761	GACAGTGTGAACTTCCCTGTAGAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGAT	2820
Qy	2783	GCCTGCTCACTGCCTCTGTTCCCTGTGTGGCTTGTGTGTGACACTCAGACTTTTGGAGGT	2842
Db	2821	GCCTGCCACCGGCTATTTCCCTGTGTGGCTGTGTGTGATACCCGGACCTTGGAGGT	2880
Qy	2843	GTCTGTGACTACTCAGGTTATGCCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAG	2902
Db	2881	GCAGAGCGACTACTTCCAGCTATGCCCGGACCTCCATCAGAGCGAGTCTCACTTCAACCG	2940
Qy	2903	TGCTCTTCAAAGCTGGGAGACCATGGGAAACAGCTCCTGTCTGGCTTTGGTGTGAAGTG	2962
Db	2941	CGGCTTTCAAGCTCGGAGGAAACATGCTCGCAAACTCTTTGGGGTTTGGGGTGAAGTG	3000
Qy	2963	TCACGGTCTATTTCTAGACTTGCAGTGAACAGCCTCCAGACAGCTCTGCATCAATATATA	3022
Db	3001	TCACAGCGTGTCTTGATTTGCAGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA	3060
Qy	3023	CAGAGTCTTCCGCTTCCAGGCTTACAGGTTCCATGCATGTGTATTCAGCTTCCCTTTGA	3082
Db	3061	CAAGATCCTCTGCTCGAGCGGTACAGGTTTCAACGATGTGTGTGTGAGCTTCCCATTTCA	3120
Qy	3083	CCAGCGTGTTAGGAAGACCTCACTTTCTTGGGCATCATCTCCAGCCAAAGCATCTCTG	3142
Db	3121	TCAGCAAGTTTGGGAAGAACCCACATTTTCTCGCGCTCATCTCTGACACGGCTCCCT	3180
Qy	3143	CTGCTATGCTATCTGAGGTCAAGATCCAGGAATGACATAAG-----GCCTC	3193
Db	3181	CTGCTACTCCATCTGAAGCCAAAGAACGACGGGATGTCTGCGGGGCGCAAGGGCGCGCG	3240
Qy	3194	TGGCTCTCTTCTCCTGAAGCCGCAATTTGGCTCTGCTTACCCAGGCTTCTGTCTCAAGCT	3253
Db	3241	CGGCCCTCTGCCCTCCGAGGCGGTGCAGTGGCTGTGCCCAACGATTTCTGTCTCAAGCT	3300
Qy	3254	GGCTGCTCATTTCTGTGATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAACACT	3313
Db	3301	GACTCGACACCGTGTCACTTACGTGCCACTCTCTGGGGTCACTCAGSACAGCCACGACGA	3360
Qy	3314	GCTGTGCGGAAGCTCCGAGGCGGACATGACCATCTTAAAGCTCAGCTGACCCAGC	3373
Db	3361	GCTGAGTCGGAGGCTCCCGGGGACGAGCTGACTGCCCTTGAGGGCGCGACGCCAACCGCG	3420
Qy	3374	CCTAAGCACAGACTTTTCAGACCAATTTTGACTAA	3407
Db	3421	ACTGCCCTCAGACTTCAAGACCACTCTCTGGACTGA	3454

RESULT 10  
US-09-733-294A-3  
; Sequence 3, Application US/09733294A  
; Patent No. 6492171  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: William Gaarde  
; APPLICANT: Susan M. Freier

APPLICANT: Edward V. Wanciewicz  
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
FILE REFERENCE: ISPH-0527

CURRENT APPLICATION NUMBER: US/09/733,294A

CURRENT FILING DATE: 2000-12-07

PRIOR APPLICATION NUMBER: 09/572,423

PRIOR FILING DATE: 2000-05-16

NUMBER OF SEQ ID NOS: 108

SEQ ID NO 3

LENGTH: 4015

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (56)...(3454)

09-733-294A-3

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

```
23 CCCGCTTGACCAATGACCCCGCTCCTCGTTGCGCCCGCGGTGCGCTCTGCTGCG 82
40 CCCGCTTGACCAATGACCCCGCTCCTCGTTGCGCCCGCGGTGCGCTCTGCTGCG 99
83 CAGCCGATACCGGAGGTGTGCGCGCTGCGCAACCTTTGTCGCGCGCTGCGGCGCCGAGGG 142
100 CAGCCACTACCGGAGGTGTGCGCGCTGCGCAACCTTTGTCGCGCGCTGCGGCGCCGAGGG 159
143 CAGCGCGCTTGTCMAACCCGGGAGACCGAGATCTACCGACTTTGGTTGCCCAATGCTT 202
160 CTGGCGGTGTGTCAGCGCGCGGACCCCGCGGCTTTCCGCGCGCTGCTGCGCGCTT 219
203 AGTGTGATGACTCGGCGCTCAGAGCTTCCACCTGCGGACCTTTCTTCTTCCACAGGTGTC 262
220 GGTGTGCTTGCTTGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 279
263 ATCCCTGAAGAGCTGTGTGCGCGCTGCGCAACCTTTGTCGCGCGCTGCGGCGCCGAGAA 322
280 CTGCTGAAGAGCTGTGTGCGCGCTGCGCAACCTTTGTCGCGCGCTGCGGCGCCGAGAA 339
323 CTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 382
340 CTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 399
383 CACTAGTACGCTGCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
400 CACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
443 TGCATGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502
460 GCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
503 CTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 562
520 CTGCGCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 579
563 GTACCAATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
580 GTACAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639
623 CCGACCGTGGCGAGGATTTCACTAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
640 GCGTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
683 TGCACGAGGACCGAACCCTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742
673 CGTACGAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
743 TCTCACCTACAGTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
733 CAGTGCACCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792
```

```
QY 803 GGAGAGGAGACCCACAGAGAGGTGCTACCAACCCCATCAGCAATCATGGTGCCAAAG 862
DB 793 GGAGCGAGACCGCGTTGGGACAGGGTCTTGGGCCCCACCCCGGACAGAGCGGTGACCGAG 852
QY 863 TCCTGTCTGGTCCCGGAGGTGCTTACT-----GCAGAGAAAGATTTGTCTTTAA 913
DB 853 TGACCGTGGTTTCTGTGTGTGTCTACCTGCCAGAGCCCGCGAAGAGCCACTCTCTTTGGA 912
QY 914 AGAAAGGTGTGACCTGAGTCTCTC---TGGGTGGTGTGCTGTAAACACAGCCCGAG 970
DB 913 GGGTGGCTCTCTGGCACCGGCACTCCACCCATCCGTGGGCGCCGACAGCCAGCCGCGG 972
QY 971 CTCACAT---CTCTGCTGTACACCCCGCAAAATGCTTTTTCAGCTCAGGCCATTTAT 1027
DB 973 CCCCCATCCACATCGGCGCCACCGTCCCTGGGACACGCTTGTCTCCCGGTGTACGC 1032
QY 1028 TGAGACAGACATTTCTTTTACTTCCAGGGAGATGGCCAGAGCGTCTTAACCCCTCAT 1087
DB 1033 CGAGACCAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
QY 1088 CTTACTCAGCAACTCTCAGCTTAATCTTGTGCGGCGAGAGACTGGTGGAGATCATCTT 1147
DB 1090 CTTACTCAGCTCTCTCAGGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
QY 1148 TCTGGGCTCAAGGCTTAGACATCAGGACCACTCTGAGGACACACCGTCTATCGCGTCG 1207
DB 1150 TCTGGTTCAGGCGCTGATGCGAGGACTCCCGCAGGTTGCGCCGCTGCGCCAGCG 1209
QY 1208 ATACTGCGATCGGCGCTGCTTCCACAGCTGCTGTTGAACCATGACAGATGCGCAATA 1267
DB 1210 CTTACTGCAATGCGGCGCTGCTTCTGAGTGTGCTTGGGACCAAGCGGCTGCGCCCTA 1269
QY 1268 TGTCAAGCTCTCAGTGTCAATTCAGGTTTGAACAGCAAAACCAACAGGTGACAGTGC 1327
DB 1270 CGGGTGTCTCTCAAGACGCTGCGCGCTGCGAGTGTGCGGTCAACCCAGCAGCGCGTGT 1329
QY 1328 CT-----TGAAACACAGCGCC 1342
DB 1330 CTGTGCGCGGAGAACCGCCAGGCTCTGTGGGCGCCCGGAGGAGGAGACACAGACCC 1389
QY 1343 ACCGACCTCATGATTTGCTCGCGCTGCACAGAGTCCCTGGCAGGTATATGTTTTCT 1402
DB 1390 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449
QY 1403 TCGGCGCTCTCTGCAAGGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1462
DB 1450 GCGGCGCTGCTGCGCGGCTGCTGCGCGCTGCGGCTGCGGCTCCAGGCTCCAGGCAAC 1509
QY 1463 CCGCTTCTTTAAGAACTTAAAGAGTTCTATCTGTTGGGAAATACGGCAAGCTATCACT 1522
DB 1510 CCGCTTCTCAGGAACACCAAGAGTTCTATCTCTGCGGAGAGCATGCAAGCTCTGCT 1569
QY 1523 GCAGGAATGATGTGGAAGATGAAGTAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
DB 1570 GCAGGAGCTGACGTGAAGATGAGCGTGCAGGCTGCGGCTGCGGCTGCGGAGGCGCG 1629
QY 1583 GAAGGACCGTGTCCCGCTGCGAGCACCGTCTGAGGGAGAGATCCTGGCTACGTTCTCT 1642
DB 1630 GGTGCTGTGTTCCCGCGCAGAGCACCGTCTGCTGAGGAGATCTCTGGCCAGTTCTCT 1689
QY 1643 GTTCTGCTGATGACACATACGTGTGATCAGTGTGTTAGGTCACTTTTATCATCACA 1702
DB 1690 GCATGCTGATGATGTGTACCTGCTGAGCTGCTCAGGTCTCTTTTATGTCTCACGGA 1749
QY 1703 GAGCACATTCAGGAAGAACAGGCTCTTCTTACCGTAAAGTGTGTGAGCAAGCTGCA 1762
DB 1750 GACCACTGTTCAAAAGAACAGGCTCTTTTCTACCGAAGATGTCTGAGCAAGTTGCA 1809
QY 1763 GAGCATTTGAGTCCAGCAACCTTGAAGAGTGTGCGCTACCGGAGCTCTCAAGAGA 1822
DB 1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGAGCTGTGGAAGCAGA 1869
QY 1823 GGTGAGGATCACCAGACACCTGGTGTAGCCATGCGCCATCTGACAGTCTGCGCTTCA 1882
```

1870	GGTCAGGCAGCATCGGGAAGCCACAGGCCCGCCCTGTGTGACGTCCAGACTCGCCTCATCCC	1929
1883	CAAGCCCAACGGCCCTCGCGGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCCTTT	1942
1930	CAAGCCTGACGGCTCGCGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTT	1989
1943	GGGCAGAGGAGCAGAGCCGCCAGCANTTTCACCCAGCGCTCAAGACTCTCTTCAGCATGCT	2002
1990	CCGCAGAGAAAGAGGGCCAGCGCTCACTCTGAGGGTGAAGGCACCTGTTTCAGCGTGCT	2049
2003	CAACTATCAGCGGCAAAAAATCCTCACCTTATGGGGTCTTCTGTACTGGGTATGAATGA	2062
2050	CAACTACAGCGGCGCGCGCCCGCCCTCTGTGGCGCCTCTGTGCTGGGCTGGACGA	2109
2063	CATCTACAGGACCTGGCGGGCTTTGTGCTGGGTGTGGTGTCTGTGACACAGACACCCAG	2122
2110	TATCCACAGGGCCCTGGCGCACCTTCGTGTGTGGTGTGGGGCCAGGACCCGCGCCCTGA	2169
2123	GATGTACTTTTGTTAAGGCAGATGTGACCCGGGGCCTATGATGTCATCCGCCAGGCTAAGCT	2182
2170	GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCCAGGACAGCCT	2229
2183	GGTGAAGTTGTTCGCAATATGATCAGGCACTCGGAGACAGTACTGTATTCGCCCAAGTA	2242
2230	CACGAGGTCTATGCCACAGCATCATCA---AACCCAGAACACACTACTCGTGGCGTGGTA	2286
2243	TGCAGTGTCCGGAGATAGCCAGGCCAAGTCCACAACTCTTTTAGGAGACAGGTCACTAC	2302
2287	TGCCGTGTCCAGAGGCCGCCCATGGGCACTGCCGANGGCCTTCAGAGCCACGTCCTC	2346
2303	CACCTCTCTGACTCCAGCCCATACATGGGCGAGTTCTTTAAGCACTTGCAGAGATTCCAGA	2362
2347	TACCTTGACAGACCTTCAGCGCTACATGCGACAGTTGTGGTGTCACTTCGACGAG-----GA	2400
2363	TGCCAGTGCACGTGAGGAACCTCCGTGTGTCATCGAGCAGAGCACTCTCATGAATGAGACG	2422
2401	GACCAGCCCGCTGAGGAGATGGCGTGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAG	2460
2423	CAGCAGCCTGTTGACTTCTTCCTGCACTTCCTCGCTCAAGTGTGCTGAAGATTGCTGA	2482
2461	CAGTGGCCTCTTCACAGCTTCTCTCACTCACTGCTTCACTGTGCCACACGCGCTGCGCATCAGGG	2520
2483	CAGGTGCTATACGACAGTGCAGGGCATCCGCCAGGGCTCCAGCCTATCCACCTGCTCTG	2542
2521	CAAGTCTCATGTCAGTGCACGGGATCCCGCAGGGTCCATCTCTCCACGCTGCTCTG	2580
2543	CAGTCTGTGTTTCGGAGACATGGAGAACAAAGCTGTGTTGTGTGAGGTGCAGCGGATGGGTT	2602
2581	CAGCCTGTGCTACCGCGCACATGGAGAACAAAGCTGTTTGGCGGGATTCCGGCGGACGCGCT	2640
2603	GCTTTTACGTTTTGTGATGACTTCTGTTTGTGAGCGCTCACTTGTGACCAAGCAAAAC	2662
2641	GCTCTCGGTTTGTGATGATTTCTTTGTTGTGACACTCACTCACCCACCGGAAAC	2700
2663	CTTCTCAGCACCTGTGTCACGCGTTCTGTAGTATGGGTGCATGAATAACTTGCAGAA	2722
2701	CTTCTCAGGACCTGTGTCGAGGTGTCCTGATATGGCTGCGTGTGAACTTGGCGAA	2760
2723	GACGTGTGAACTTCCCTGTGAGCCTGTACCTCGGTGTGCGAGCTCCATACCACT	2782
2761	GACGTGTGAACTTCCCTGTGAGAGCAGAGGCCCTGTGGGTGSCACGGCTTTTGTTCAGAT	2820
2783	GCTCTCACTGCTGTTTCCCTGTGTGTGCTTGTGCTGGACACTCAGACTTTGGAGGT	2842
2821	GCGGCCCAACGGCCTATTCCCTGTGTGGGCTGTCTGTGGATACCCGACCCCTGGAGGT	2880
2843	GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAGACAGGCTCAGCTTCCAGAG	2902
2881	GCAGAGGCACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCATCTTCAACCG	2940
2903	TGCTTTCAAAGCTCGGGAAGACCATTCGGGAACAAGCTCTCTGTGCTTGTGGGTTGGAAGTG	2962

2941	Db	CGGCTTCAAGCGCTGGGAGGAACATGCGTCGCGAAA	CTCTTTGGGGGTCTTGGCGCTGAAGTG	3000
2963	Qy	TCAGGGCTATATTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGCTCTGCATCAATATATA		3022
3001	Db	TCACAGCGCTGTTCTCGGATTTGAGGTGAACAGCCTCCAGACGGGTGSCACCAACATCTA	3060	
3023	Qy	CAGGATCTTCTCGCTTCAGGCTTACAGGTTCCATGCATGTGTGATGTGATTCAGCTTCCCTTTGA	3082	
3061	Db	CAGGATCCTCTCTGCTCGAGCGGTACAGGTTTCACGCATGTGTGCTGTGACGTCTCCCATTTCA	3120	
3083	Qy	CCAGCGTGTAGGAAGAACCTCACA	TTCTTTCTGGGCGATCATCTCCAGCCAAAGCATCCTG	3142
3121	Db	TCGACAACTTTGGAAGAAACCCACA	TTTTCCTGCGCTCATCTCTGACACGGCTCCCT	3180
3143	Qy	CTGCTATGCTATCCTGAAAGTCAAGAAATCCAGGATGACACTAAG	-----GGCTC	3193
3181	Db	CTGTACTCCATCCTGAAAGCCAGAACCCAGGATGTGCTGGGGGCCAAGGGCGCGC	3240	
3194	Qy	TGGCTCCTTTCTCCTCGAAGCCGCAATTTGGGCTCTGTACACAGGCTTCTCTGCTCAAGCT	3253	
3241	Db	CGGCCCTCTGCGCTCCGAGGCGGTGCASTGGCTGTGCCACAGCAATTCCTGTCTCAAGCT	3300	
3254	Qy	GGCTGCTCATTTCTGTGATCTACAATATGCTCTCTGGGA	CCTCTGTAGGAACAGCCCAAAACT	3313
3301	Db	GACTCGACACCGTGTCACTTACGTGTGCCACTCTCTGGGGTCACTCAGGAACAGCCGACGCA	3360	
3314	Qy	GCTGTGCGGAAGCTCCAGAGCGGCAATATGACCAT	CTTTAAAGCTGCAGCTGACCCAGC	3373
3361	Db	GCTGAGTCGGAAGCTCCCGGGGACGAGCTGACTGCCCTTGAGGCGCGGACCAACCCGGC	3420	
3374	Qy	CCTAAGCAACAGACTTTCAGACCAATTTTGACTAA	3407	
3421	Db	ACTGCCCTCAGACTTCAAGACCACTCTGTGACTGA	3454	

RESULT 11  
US-08-974-549A-343  
; Sequence 343, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joschim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Moxin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 343:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4037 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /note= "refined sequence of hTERT cDNA"  
-08-974-549A-343  
Query Match 45.7%; Score 1596; DB 3; Length 4037;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 2376; Conservative 0; Mismatches 970; Indels 108; Gaps 9;  
23 CCGGGCTTTGACCAATGACCCGGCTCCCTGTTGGCCCGCGTGGCGCTCCCTGCGCTCCCTGCTGG 82  
40 CCGGGCCACCCCGCGATGCGCGCGCTCCCGCTGGCGAGCCGCTCCCTGCTGG 99  
83 CAGCGGATACCGGAGGTGTGGCCGCTGGCAACCTTTGTGGCGCGCTGGCGCCGAGG 142  
100 CAGCCACTACCGGAGGTGTGGCCGCTGGCAACCTTTGTGGCGCGCTGGCGCCGAGG 159  
143 CAGCGCGCTTTGTGCAACCCGGGACCCGAGATCTACCGCACTTTGGTTGCCCAATGCC 202  
160 CTGGCGCTGGTGCACCGCGGGAACCGCGGCTTTCCGCGCGCTGGTGGCGCGAGTCC 219  
203 AGTGTGCATGCACTGGGGTGCACCGCTTCACCTGCGCACTTTCCCTTCAACGAGGTGC 262  
220 GGTGTGCGTGGCTGGGAGCGACCGCGCCCGCCCGCGCCCTCCCTTCCGCGAGGTGC 279  
263 ATCCCTGAAGAGCTGGTGGCCAGGTTGTGAGAGCTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCGCTGAAGAGCTGGTGGCCGAGTGTGAGAGGCTGTGAGGCGCGCGCGAGAA 339  
323 CGTGTGGCTTTGGCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 382  
340 CGTGTGGCTTTGGCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 399  
383 CACTAGTACGCTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 442  
400 CACCACCGCTGGCGAGTACTCTGCGCGGCTGTGGTGGCCCGAGGCTCTGGGCTCCAGGCA 459

QY 443 TGCATGGATGCTACTGTTGAGCGAGTGGCGACGACCTGTGCTTACCTGCTGGCACA 502  
DB 460 GGGGTGGGGCTGCTGCTGCGCGCGTGGCGACGACGCTGCTGGTTCACTGCTGGCAG 519  
QY 503 CTGTGCTCTTTATCTTCTGTTGCGCCCGACAGCTGTGCTACCAAGTGTGTGGGTCTCCCT 562  
DB 520 CTGGCGCTCTTTGCTGCTGGTCTCCAGCTTGGCGCTACCAAGTGTGTGGCGCGCGCT 579  
QY 563 GTACCAAAATTTGTGCGACGAGATATCTGGCGCTTGTGCTCCCTAGTTCAGGCGCCAC 622  
DB 580 GTACCAAGCTCGGCGCTGCCACTCAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 639  
QY 623 CCGACCGGTGGCGAGGAATTTCACTAACTTAGTCTTACCAAGATCAAGAGCAGTAG 682  
DB 640 GCGTCTGGGATGC-----GACCGGCTTGAACCATAG 672  
QY 683 TCGCCAGGAAGCACCGAAACCTGCTGCTTGGCATCTCGAGGTACAAGAGGATCTGAG 742  
DB 673 CGTCAGGGAGCGCGGGTCTCCCTGCGCGCTGCGAGCGCGGGTGGAGAGCGCGGGG 732  
QY 743 TCTCACCAGTACAGTGTGCTTCACTAAGAGGCGAGATCTATCTCTCTCCGAGAGT 802  
DB 733 CAGTGCCAGCGGAGTCTGCGCTTGGCCAGAGCGCGAGCGTGGCGCTGCGCTCTGAGCC 792  
QY 803 GGAGGAGGAGCGCCACAGGCGAGTGTCTACCAACCCCATCAGGCAATCATTGGTGGCCAG 862  
DB 793 GGAGCGGACCGCGCTTGGCGAGGGTCTTGGGCGCACCGCGCGAGCGCGTGGACCGAG 852  
QY 863 TCCTGCTCGTCCCGCGAGTGTCTACT-----GCAGAGAGATTTCTCTCTAA 913  
DB 853 TGACCGTGGTCTTGTGTGTGTGTCACCTGCGAGACCGCGCGAGGAGCACTCTTTGGA 912  
QY 914 AGGAAAGTGTCTGACCTGAGTCTCTC---TGGGTGCGTGTGCTGTAAACACAAGCCAG 970  
DB 913 GGGTGGCTCTCTGGCAGCGCGCACTCCACCATCGTGGCGCGCGAGCACCAAGCGGG 972  
QY 971 CTCACAT---CTCTGCTGTACACACCGCGCGCAAAATGCTTTTCTGAGTCTAGGCTAT 1027  
DB 973 CCCCCCATCCACATCGCGCGCACCGTCCCTGGGACAGCGCTTGTCCCCCGGTGTACGC 1032  
QY 1028 TGAGCAGACATTTCTTTTACTCCAGGGAGATGGCCAGAGCGTCTAAACCCCTCAT 1087  
DB 1033 CGAGACGACATCT 1089  
QY 1088 CTTACTCAGAACTCTCAGCTAACTGAGTGGGGCGAGAGACTGGTGGAGATCATCTT 1147  
DB 1090 CTTACTCAGCTCTCTGAGGCGCGAGCTGACTGGCGCTCGGAGGCTCTGGAGACCATCTT 1149  
QY 1148 TCTGGGCTCAGGCTTAGGACATCAGGACCATCTCGAGGACACACCGTCTATCGGCTCG 1207  
DB 1150 TCTGGGTTCCAGGCTTGGATCGAGGACTCCCGAGGTTGCCCGCTGCGCGCGAG 1209  
QY 1208 ATACTGCGAGATGCGCGCTCTGTTCAACAGCTGTGTGTAACCATGCGAGAGTGCAATA 1267  
DB 1210 CTACTGCAATGCGCGCTCTGTTCTGAGAGTGTCTGGGAAACCAAGCGCGAGTGGCCCTA 1269  
QY 1268 TGTGAGACTCTCAGTCACTTGCAGGTTTGAACAGCAAAACCAAGAGTGCAGATGC 1327  
DB 1270 CGGGGTGCTCTCTCAAGACGCACTGCGCGCTGCGAGTGTGCGGTCAACCGAGCGCGGT 1329  
QY 1328 CT-----TGAAACACGAGCC 1342  
DB 1330 CTGTGCGCGGAGAGCGCGCGCTGTGGCGCGCGCGCGAGGAGGAGGACAGAGCC 1389  
QY 1343 ACCGAGCTCATGATTTGCTCGCGCTGCGAGCTGCTGCGAGGATATGTTTCT 1402  
DB 1390 CCGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
QY 1403 TCGGCGCTGTCTCTGCAAGTGTGTCTGCTAGTCTCTGGGTACCGAGGCAAAATGAGCG 1462  
DB 1450 CGGGGCTGTCTGCTGCGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509

1463	CGCTTCTTTAAGAACTTAAAGAAATTCATCTCGTTGGGGAATACGGCAAGCTATACAT	1521
1510	CGCTTCTCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAAGCTCTGGCT	1569
1523	GCAGGAACCTGATGTGGGAAGATGAAGTAGAGGATGCACTGGCTCCGCGAGCCCGGG	1582
1570	GCAGGAGCTGACGTGGAGATGACGCTGCGGACTGGCTTGGCTTGGCTGCGAGGACCCAGG	1629
1583	GAAGGACCGTGTCCTCCGCTGCGAGACACCGTCTTGAAGGAGAGAGATCTTGGCTAGCTTCCT	1642
1630	GGTGGCTGTGTTCGGCCGCGAGACACCGTCTGCGCTGAGAGAGATCTTGGCCAAAGTTCCT	1689
1643	GTCTGGCTGTATGGACACATACGCTGGTACAGCTGTTAGTCACTTCTTTTATCATCACAGA	1702
1690	GCATCTGGCTGATGAGTGTACGCTGTGAGCTGCTCAGGCTCTTCTTTTATGTACAGGA	1749
1703	GAGCACTTCCAGAGAACAGGCTCTTCTTCTACCGCTTAAGAGTGTGTGAGCAAGCTGCA	1762
1750	GACCACGTTTCAAGAAGACAGGCTCTTTTCTACCGCGCAGTGTCTGAGCAAGTTGCA	1809
1763	GAGCACTGGAGTCAAGGCAACACCTTGGAGAGTGGCGCTACGGGAGCTGTCACAAGGGA	1822
1810	AAGCAATTGGAATTCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTTCGGAAGCAGA	1869
1823	GCTCAGGCATCACAGGACACCTTGGCTAGCCATCCCATCTCCAGACTCGGCTTCATCCC	1882
1870	GCTCAGGCAGCATCGGGAAGCCAGGCCGCCCTGCTGAGCTCCAGACTCCGTTTCATCCC	1929
1883	CAAGCCCAACGGCTCGGCGCCATTGTGAACATGAGTTATACATGGGTACACAGAGTTT	1942
1930	CAAGCTTCACTGCGGGCTGCGGCGGATGTGTGAACATGTGACTACGCTGTGGAGGCGAGAAGCTT	1989
1943	GGCGAAGAGGAGCAGGCCACGACTTTCACCCAGCGCTCTCAAGACTCTCTTCAGCATGCT	2002
1990	CGCAGAGNAAGAGGGCCGAGCGTCTCACTCGAGGGTGAAGGCACCTGTTACGCGTGCT	2049
2003	CAACTATGAGCGGGAJAAAAATCTCTACCTTATGGGCTCTTCTGTACTGGGTATGAATGA	2062
2050	CAACTACAGCGGCGCGGGCGCCCGGCTCTCTGGCGCGCTCTGTGCTGGGCTGTGACGA	2109
2063	CATCTACAGGACCTTGGCGGGCTTTGTGCTGGGTGGTGGTCTGTGAGACAGACACCCAG	2122
2110	TATTCACAGGCGCTTGGCGACCTCTGTGCTGGTGTGGGCGGCTCTGTGCTGGGCTGTGACGA	2169
2123	GATGTATTTTGTAAAGCAGATGTGACCGGGGCTATGATGTCATCCATCCCGCAGGTAAGCT	2182
2170	GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGCTACGACACCATCCCCCAGGACAGCT	2229
2183	GGTGAGGTTCTTGCCATATGATCAGGCACTCGGAGAGCAGTACTATCTCCGCAAGTA	2242
2230	CACGAGGTCTATGCGCAGCATATCA---AAGCCAGAAACAGTACTCGTGGCTCGGTA	2286
2243	TGCAGTGTCCGAGAGATAGCCAAAGGCCAAAGTCCCAAGTCCCTTTAGGACACAGGTACAC	2302
2287	TGCCGTGTTCAGAAAGCCGCCATGGGCACTGTCGCAAGGCTTCAGAGCCACTTCACAGCCAGTCTC	2346
2303	CACCTCTCTGACCTTCAGCGATACATGGGCGAGTTCCTTAAAGATCTGTGAGAGATTCAGA	2362
2347	TACCTTGAACAGACCTTCAGCCGCTACATGCGACAGTTTGTGGTCTCACTCTGCAG-----GA	2400
2363	TGCCAGTGCATCGAGGAACCTCCGTTGTTCATCGAGCAGACATCTCTATGAATGAGAGCAG	2422
2401	GACCAAGCCGCTGAGGATGCGCTGTCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG	2460
2423	CAGCAGCTGTTTGACTTCTTCTGCACTTCTCGGTCACTCGGTGTGCTGAAGATTTGGTGA	2482
2461	CAGTGGCTCTTTCGACGCTCTTCTCACTGCTTCATGTGCCACACAGCCGCTGCGCATCAGGG	2520
2483	CAGGTGCTATACGAGTGCAGGGCATCCCCAGGGCTCCAGCTTATCCACCTGCTCTG	2542
2521	CAAGTCTCTACGTTCAGTGCACGGGATTCGCCAGGGGCTCCATCTCTCTCACTGCTCTG	2580
2543	CAGTCTGTGTTTCGGAGACATGAGAAACAAGCTGTTTGTGAGGTGCAGCGGGATGGGTT	2602

2581	Db		CAGCCTGTGTCACGGGACATGGAGAACAGCTGTGTTGGGGATTTCGGCGGACGGGCT	2640
2603	Qy		GCTTTTACGTTTTTGTTGATGACTTCTGTGTGTGAGCGCTCACTTGGACCAAGCAAAAC	2662
2641	Db		GCTCCTGCGTTTGGTGGATGATTTCTTGTGTGTGACACTCACTCACCACCGGAAC	2700
2663	Qy		CTTCTCAGACACCTGGTTCATGGCGTTCTGATGTGGGTGATGATAACTTGAGAA	2722
2701	Db		CTTCTCTCAGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGAATTCGGAA	2760
2723	Qy		GACAGTGGTGAACCTTCCCTGTGGAGCTGGTACCTGGGTGGTGGAGCTCCATACCAGCT	2782
2761	Db		GACAGTGTGAACCTTCCCTGTGAGACGAGGCCCTGGGTGGACGGCTTTGTTGAGAT	2820
2783	Qy		GCTGTCTCACTGCGTGTTCCTCTGGTGTGGCTTGTCTGTGGACACTCAGACTTTGAGGT	2842
2821	Db		GCCGGGCCACGGGCTTATTCCTCTGGTGGCGGCTTGTCTGTGGATACCGGACCTCGAGGT	2880
2843	Qy		GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAAGCAGAGCTCACTTCCAGAG	2902
2881	Db		GCAGAGGACTACTCCAGCTATGCCCCGACCTCCATCAGAGCCAGTGTCACTTCAACCG	2940
2903	Qy		TGTCTTCAAAGCTGGGAAGACATGCGGAAACAGACTCCTGTCTGGTCTGGGTTGAAGTG	2962
2941	Db		CGGCTTCAAGGCTGGGAGGAACATGGGTGCGAACTCTTTGGGGTCTTCGGGCTGAAGTG	3000
2963	Qy		TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATA	3022
3001	Db		TCACAGCTGTTTCTGATTTGCAAGTGAACAGCTCCAGACGGTGTGCACCAACATCTA	3060
3023	Qy		CAAGATCTTCTGCTTCAGGCTTACAGGTTCCATGATGTGTGATTACGTTTCCCTTTGA	3082
3061	Db		CAAGATCTCTCTCTCGAGCGTACAGGTTTACGATGTGTGCTCAGCTCCCATTTCA	3120
3083	Qy		CCAGCGTGTTAGGAAGACCTCAATTTCTTTGGGCATCATCTCCAGCGACAGCATCTCTG	3142
3121	Db		TCAGCAAGTTTGAAGAACCCACATTTTTCTGCGGTCATCTCTGACAGGCTTCCCT	3180
3143	Qy		CTGCTATGCTATCCTGAAGGTCAAGAAATCAGGAATGACACTAAAG-----GCCCTC	3193
3181	Db		CTGCTACTCCATCCTGAAAGCCAAAGAACCGAGGGATGTCTGCTGGGGGCCAAGGGCGCCG	3240
3194	Qy		TGCTCTTTTCCCTCTGAAGCCGACATTTGGTCTCTGTACAGGCCCTTCTGCTCAAGCT	3253
3241	Db		CGGCTCTGCCCTCGAGGCGGTGAGTGGTGTGCTGCCACCAAGCAATTCCTGCTCAAGTT	3300
3254	Qy		GGTGTGCTATTTGTGTCATCTACAAATGTCTCTCTGGGACCTCTTGAGGACAGCCCAAACT	3313
3301	Db		GACTCGACACCGGTCACTTGTGTCACCTCTCTGGGGTCACTCAGGACAGCCGACAGCA	3360
3314	Qy		GCTGTGCGGAAGCTCCAGAGGCGCAATGATCATCTTAAAGCTGCAGCTGACCCAGC	3373
3361	Db		GCTGAGTCGGAAGCTCCCGGGAGCAGCGCTGACTGCCCTTGGAGCGCGAGCCNA	3420
3374	Qy		CCTAAGCACAGACTTTACAGACCATTTTGGACTAA	3407
3421	Db		ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA	3454

RESULT 12

RESOL 12  
US-08-851-843A-173  
; Sequence 173. Application US/08851843A

; Patent No. 6093809

GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H  
TITLE OF INVENTION: No 603938



NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 173:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4029 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY:  
LOCATION: 1..4029  
OTHER INFORMATION: /note= "preliminary sequence for  
human TRT cDNA insert of  
OTHER INFORMATION: plasmid pGRN121"

-08-851-843A-173

Query Match 43.5%; Score 1520.2; DB 3; Length 4029;  
Best Local Similarity 57.2%; Pred. No. 0;  
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;

23 CCGGCTTTAGACAAATGACCGGCTCTCGTGGCCCGGCTCTCGTGGCG 82  
40 CCGGGACACCCCGGATCGCGGCTCCCCGCTCCGAGCGGCTCCCTGCTGG 99  
83 CAGCCATACCGGAGGTGGCGGTGGCACTTTGTGGGCGCTGGGGCCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTCGCGGTGGCCACGTTGTGGGCGCTGGGGCCCGAGGG 159  
143 CAGCGGCTTGTCAACCGGGGACCGGAGATCTACCGACTTGTGGTGGCAATGCT 202  
160 CTGGCGCTGGTTCAGCGGGGACCGGCGCTTTCCGCGGNTGGTGGCCANTGCT 219  
203 AGTGTGATCACTGCGGTTCACAGCTCCACTGCGGACCTTTCTTCCACAGGTGTC 262  
220 GGTGTGGTCCCTGGGANGNANGGNGCCCGCCCGCCCTCTTCCGCGAGGTGTC 279  
263 ATCCCTGAAGAGCTGGTGGCCAGGTTGTGAGAGACTTGTGCGAGCGCAACGAGAGAA 322

Db 280 CTGCTCANGANCTGGTGGCCGAGTGTGTCANRANGTGTGCGANCGCGCGCAANAA 339  
Qy 323 CGTCTCGCTTTTGGCTTTGAGTGTCTTAACGAGGCCAGAGGGGCTCTCCATGCGCTT 382  
Db 340 CGTCTGGCTTTCGGCTTTCGCTGTGGACGGGGCCCGCGGGGGCCCCCGAGGCGCTT 399  
Qy 383 CACTAGTAGGCTGGTAGCTTGTGCCAACACTGTATTGTGAGACCTCGCTGTGCTGAGTGG 442  
Db 400 CACCACGAGGTGGCGAGCTTACTGTCACACAGCTGACCGACGACTGCGGGGAGCGG 459  
Qy 443 TGCATGATGCTACTGTTAGCGGAGTGGGCGACAGCTGTGCTTACTGCTGTGGCACA 502  
Db 460 GCGTGGGGCTGCTGCTGGCGCGCTGGGCGACGCTGCTGCTTCACTGCTGGCAGC 519  
Qy 503 CTGTGCTCTTTATCTTCTGTGTCGCCCGCTGCTACACAGGTGTGTGGGTCTCCCT 562  
Db 520 CTGCGGNTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
Qy 563 GTACCAATTTGTGCCACACGAGTATCTGGCCCTCTGTGTCGCTAGTTTACAGGCCAC 622  
Db 580 GTACCGCTGGCGCTGCNACTCAGGCCCGCGCCCGCCACACGCTANT----- 628  
Qy 623 CCGACCGTGGCGAGGAATTTCTACTTACCTTAGGTTCTTACACAGATCAAGAGCAGTAG 682  
Db 629 -----GGACCCGAANGCGTCTGGGATCCACGGGCTGGAAACCATAG 670  
Qy 683 TCGCAGGAAGCAGCAACCCCTTGGCTTGCCTCTCGAGGTACAAAGAGCATCTGAG 742  
Db 671 -CGTCAGGGAGCGCGGCTCCCTTGGCTGTCAGCCCGGCTGCGAGGAGCGCGGGG 729  
Qy 743 TCTCAGGATACAGTGTGCTTCACTAGAGAGGCGAGATGCTATCTGCTCCGAGAGT 802  
Db 730 CAGTGCCAGCGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789  
Qy 803 GGAGGAGGAGCGCCACAGGCGAGGTCTACCAACCCCATCAGGCAAAATCATGGTGGCAAG 862  
Db 790 GGAGGAGGAGCGCGCTTGGGCGAGGGTCTCTGGGCGCCACCGCGGCGAGCGCTGGACCGAG 849  
Qy 863 TCCTGCTCGTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 913  
Db 850 TGACCGTGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909  
Qy 914 AGGAAAGGTGCTGACCTGAGTCTCTCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973  
Db 910 GGGTGGCTCTCTGGCAGCGCT 969  
Qy 974 CACATCTCTGCTGCTCACCACCCCGCCAAATGCTTTTTCAGCTCAGGCCATTTATTGAGAC 1033  
Db 970 CCCCCCATCCACATCGCGGCGCACCGCTCTCTGCGACACGCTTGTCCCCGGGTGACGCG 1029  
Qy 1034 CAGCATTTCTTTTACTCCAGGGGAGATGCCAGAGCGCTTAAACCCCTCATTTCTCTACT 1093  
Db 1030 GAGACCAAGCACTTCTCTTACTCTCTGAGGCGACAGNACACTGCGNCCCTCTCTCTCTCT 1089  
Qy 1094 CAGCA-ACCTCCAGCTAACTTGACTGGGGCCAGAGACTGGTGGAGATCATCTTTCTGG 1152  
Db 1090 CAATATATCTCAGGCCCGAGCTGACTGCGCTTGGGAGGTTCTGAGACANTCTTTCTG 1149  
Qy 1153 GCTCAGGCTTAGACATCAGGACCTCTGAGGACACACCGCTCTATCGCTCGATCT 1212  
Db 1150 GTTCAGGCTTGGATGCCAGGA-TTCCCCGAGGTTGCCCCCGCTGCCCCAGCGTACT 1208  
Qy 1213 GCGAGATCGGCGCTCTGTTCCACAGCTGTGTGTAACCATGCGAGAGTGCCTAATATGTCA 1272  
Db 1209 GGCANATCGGCGCTCTGTTCTGAGTGTGTTGGAAACACCGCGAGTGCCTCTACGGG 1268  
Qy 1273 GACTCTCAGGTCACTTGGAGGTTTGAACAGCAACCAACAGGTGACAGATGCT--- 1329  
Db 1269 TGTTCCTCAGAGCGACTGCGCTGCGAGTGGGTTCACTCCAGGAGCGCGGTGTCTGTG 1328  
Qy 1330 -----TGAACACCGACCG 1346  
Db 1329 CCGGGAGAGCCCGAGGCTCTGTGGGCGCCCGGAGGAGGAGGAGACACAGACCCCGCT 1388

1347 CACCTCATGATTTGCTCCGCTGCAGACAGTCCCTGGCAGTATATGGTTTCTTGG 1406  
1389 CGCTGTGTGAGCTGCTCCGCCAGACAGACAGCCCTGGCAGTGTACGGCTTCTGGGG 1448  
1407 GCCTGTCTCTCAAGTGGTGTCTGCTAGTCTCTGGGGTACCAAGCACAATGAGCGCCG 1466  
1449 GCCTGCTGCGCCGGCTGGTCCGCCAGGCTCTGGGGCTCCAGGCAACGAAACGCCG 1508  
1467 TTCTTTAAGAACTTAAGAGTTCATCTCTGTTGGGAAATACGGGAAGTATCATCTGCG 1526  
1509 TTCTCTAGGAAACACCAAGATTCATCTCTCCCTGGGAAAGCATCCAAAGTCTCTGCTG 1568  
1527 GAACTGATGTGGAAGATCAAGTAGAGATTGCCACTGGCTCCGACAGACCCCGGGAG 1586  
1569 GAGCTGAGCTGGAAGATGAGCTGCGGGAATGGCTTGGCTGCGAGGAGCCAGGGGTT 1628  
1587 GACCGTGTCCCGCTGCAGACACCGCTGTGAGGAGAGGATCTCTGGCTAGTTCCTGTT 1646  
1629 GGCTGTGTTCCGGCCGACAGACACCGCTGTGCTGAGGAGATCTCTGGCAAGTTCCTG 1688  
1647 TGGCTGATGACATACAGTGTACAGTGTACAGTGTCTAGTCACTCTTTTACATCA 1706  
1689 TGGCTGATGAGTGTACGCTGTGAGCTGTCTAGGCTCTCTCTTTTATGTACGGAG 1748  
1707 ACATTTCCAGAAAGACAGCTCTCTTCTTACCGTAAGAGTGTGTGAGCAAGCTGCAG 1766  
1749 ACGTTTCAAAAGAACAGCTCTTTTCTACCGGAAGATGTCTGGAGCAAGTTGCAAG 1808  
1767 ATTGGAGTCAGCAACACCTTGAGAGTGGCTACGGAGTGTGTCAAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGCACTTGAAGAGGTGCAGCTGGGGAGCTGTGGAAGCAGAG 1868  
1827 AGGCATCACACAGACACCTGGCTAGCATGCCATCTGCAGAGTGGGCTTCATCCCA 1886  
1869 AGGCAGCATCGGAAGCAGGCCCGGCTGTGTGAGCTGCAGACTCCGCTTCATCCCA 1928  
1887 CCCAAGCCTCGGGCCCATTTGGAACATGAGTTATAGCATGGGTACAGAGTTTGG 1946  
1929 CCGTGAAGGCTCGGGCCGCTTGAAGATGAGCTTACGCTGTGGAGCCAGAGCTTCC 1988  
1947 AAGAGAGAGAGGCCAGCATTTCCAGCAGCTCTCAAGACTCTTTCAGCATGTCTCA 2006  
1989 AGAAGAAAGAGGCCGAGCTCTCACTCAGAGGTGAGGCACTGTTCAGCGTGTCA 2048  
2007 TATGAGCGCAAAACATCTCACTTATGGGTCTTCTGTACTGGGTATGAATGATC 2066  
2049 TACGAGCGGGCGCGGCCCGGCTCTGGGGCCCTCTGTGGGGCTTGACGATATC 2108  
2067 TACAGGACTCGGGGCTTTGTGCTGGCTGTGGCTCTTGACAGACACCCAGGATG 2126  
2109 CACAGGGCTGGGACCTTCTGCTGGCTGTGGGGCCAGGACCGCGCTGAGCTG 2168  
2127 TACTTTGTTAAGGAGATGTGACCGGGCCCTATGATGCCATCCCGAGGGTAACTGG 2186  
2169 TACTTTGTCAAGTGGATGTGACGGCGGTACGACACCATCCCCCAGGACAGGCTCA 2228  
2187 GAGGTGTTGCAATATGATCAGGCACTCGAGAGCAGTACTGTATCCGCCAGTATGA 2246  
2229 GAGGTGATCGGCAGCATATCA---AACCACAGAACACGTAAGTGGTGGTATGCC 2285  
2247 GTGGTCCGAGAGATAGCAAGGCCAAGTCCACAAGTCTTTTAGGAGAGAGGTCA 2306  
2286 GTGGTCCAGAGAGGCCGCCATGGGACGTCCGCAAGGCTTTCAAGAGCCACGCTCT 2345  
2307 CTCTCTGACCTCCAGCCATATATGGGCCAGTTCCTTAAGCATCTGAGGATTCAG 2366  
2346 TTGACAGACCTCCAGCCGCTATATCGGACAGTGTGTGGCTCACCTGCGAG----- 2399  
2367 AGTGCATGAGGAATCCGTTGTATCGAGAGAGCATCTTATGATGAGAGCAGCAG 2426  
2400 AGCCGCTGAGGGATGCGCTGCTATCGAGCAGAGCTCTCCCTGATGAGGCCAGCT 2459

Qy 2427 AGCCTTTTTCATCTTCTTCTGCACTTCTGCTGCTACAGTGTCTTAAGATTGGTGACAGG 2486  
Db 2460 GGCTCTTTCGAGCTTCTTCTACGCTTCTATGCTCCACACGCGCTGGCATCAGGGGCAAG 2519  
Qy 2487 TGTATACGCAAGTGCAGGGCATCCGCCAGGCTCCAGCCTATCCACCCCTGCTCTGCACT 2546  
Db 2520 TCTACGTCAGTGCAGGGATCCCGAGGGCTCATCTCTCCACGCTGCTCTGCAAG 2579  
Qy 2547 CTGTGTTCCGAGACATGAGAAACAAGCTGTTGCTGAGGTGCAGCGGATGGGTGCTT 2606  
Db 2580 CTGTGTTCCGAGACATGAGAAACAAGCTGTTGCGGGATTCGCGGGAGCGGGTGTCT 2639  
Qy 2607 TTAGCTTTTGTGATGATCTTCTGTTGTTGAGCGCTCACTTGAACCAAGCAAAACCTTC 2666  
Db 2640 CTGGTTTGGTGGATGATTTCTTGTGTTGAGCACTCACTCAACCCACGCAAACTTC 2699  
Qy 2667 CTAGACACCTGTGCTCATGGCTTCTGAGTATGGTGTGATGATGATAAACTTGCAGAGACA 2726  
Db 2700 CTAGACACCTGTGCTCATGGCTTCTGAGTATGGTGTGCTGCTGAGTGTGCGGAAGACA 2759  
Qy 2727 GTGGTGAACCTTCCCTGTGAGCTGTGTACCTCCCTGGGTGTGAGCTCCATACAGCTGCT 2786  
Db 2760 GTGGTGAACCTTCCCTGTGAGAGAGGCGCTTGGTGTGACCGCTTTGTTGAGATGCG 2819  
Qy 2787 GCTCACTGCTGTTTCCCTGGTGTGCTTGTGCTGTGACACTCAGACTTTGGAGGTTC 2846  
Db 2820 GCCACGGCCCTATTCCTCCCTGGTGGCGCTGTGCTGTGATACCGGACCTGGAGTGCAG 2879  
Qy 2847 TGTGACTACTCAGTGTATGCCAGACCTCAATTAAGAGAGCTCACCTTCCAGAGTTC 2906  
Db 2880 AGGACTACTTCCAGTGTATGCCAGACCTCCATCAGAGCAGTCTCACCTTCAACCGCGC 2939  
Qy 2907 TTCAAAGCTGGGAAGACCATGCGGAACAAGCTCTCTGCTGCTTCCGTTTGAAGTGTCA 2966  
Db 2940 TTCAAAGCTGGGAAGACCATGCTGCGCAACTCTTTGGGCTTTCGGCTGAAGTGTCA 2999  
Qy 2967 GGTCTATTTCTAGACTTTCAGGTGAGACAGCTTCCAGAGCTGATGATCAATATATACA 3026  
Db 3000 AGCCTGTTTCTGATTTGAGGTGAGACAGCTTCCAGCGGTGTGACCAACATCTACAG 3059  
Qy 3027 ATCTTCTGCTTCCAGGCTTACAGTTCATGATGATGATGATGATGATGATGATGATG 3086  
Db 3060 ATCTTCTGCTGAGGCTTACAGTTCATGATGATGATGATGATGATGATGATGATGAT 3119  
Qy 3087 CGTGTAGGAAGAACTCTCATCTTTCTGGGCTATCTTCCAGCCAAAGCATCTCTGCTGC 3146  
Db 3120 CAAGTTTGAAGAAACCCCACTTTTCTGCTGCTGCTCTGATGACAGGCTCTCTCTGCT 3179  
Qy 3147 TATGCTATCTGAGGTCAAGATCCAGGAATGACACTAAAG-----GCCCTCTGGC 3197  
Db 3180 TACTCATCTGAAAGCCAGAAAGCAGGAGATGCTGCTGGGGCCAGGGCGCCCGCGC 3239  
Qy 3198 TCCTTTCTCTGAAGCCGACATTTGGTCTCTTACAGGCTTCTCTCTCAAGTGGCT 3257  
Db 3240 CCTCTGCTTCCAGGCGCTGCACTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3299  
Qy 3258 GCTCATCTGCTATCTACAAATGCTCTGCTGGAACCTTGAAGACAGCCCAAACTGCTG 3317  
Db 3300 CGACACGCTGCTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3359  
Qy 3318 TGCGGAAGCTTCCAGAGGCGCAATGACCATCTTTAAAGTGTGAGCTGAGCCAGCCCTA 3377  
Db 3360 AGTCGGAAGCTTCCCGGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3419  
Qy 3378 AGCAGACATTTTTCAGACCAATTTTGGACTMA 3407  
Db 3420 CCCTCAGACTTCAAGACCATCTCTGACTGA 3449

RESULT 13

US-08-974-549A-292

; Sequence 292, Application US/08974549A

; Patent No. 6166178

```

; OTHER INFORMATION: /note= "preliminary sequence for
; OTHER INFORMATION: human TRT cDNA insert of
; OTHER INFORMATION: plasmid pGRN121"
; US-08-974-549A-292

Query Match          43.5%; Score 1520.2; DB 3; Length 4029;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;

QY 23 CCCGGCTTGAGACACAATACCCGCGCTCTCGTTGCCGCCGCGTGCCTCTCTGCTGCG 82
Db 40 CCCGGCACCCCGCGATGCGCGCGCTCCCGCTCCGAGCGGTGCGCTCCCTGCTGCG 99
QY 83 CAGCGGATACCGGAGGTGGCGCTGCAACCTTTGTGCGCGCTTGGCGCCCGAGG 142
Db 100 CAGCCACTACCGGAGGTGTCGCCGTGSCACGTTGTGCGCGCTTGGGCGCCAGG 159
QY 143 CAGCGGCTTGTCACCCCGGACCCGAGATCTACCGCACTTTGTGTGCCAAATGCT 202
Db 160 CTGCGGCTGTGTCAGCGCGGGACCCGCGGCTTTCGCGCGTGTGGCCANTGCT 219
QY 203 AGTGTGATGACACTTGGGGCTCAGGCTCCACTCGCGACCTTTCCTTCCACGAGGTGC 262
Db 220 GGTGTGCGTGGCCCTGGGANGNNGGCCCGCCGCCCTCTCTTCCGCGAGGTGC 279
QY 263 ATCCCTGAAGAGCTGTGTGCCAGGGTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322
Db 280 CTGCTGAANGANCTGGTGCCCGAGTGCTGCANANGCTGTGCANCGCGCGGGAANA 339
QY 323 GGTGCTGCTTTGGCTTTGAGCTGCTTAACGAGGCCAGAGCGGGCTCCCATGGCTT 382
Db 340 GGTGCTGCTTTGCTTTCGCTTTCGCTGTGACGCGGCGCCCGCGGGCGCCCTCGAGGCTT 399
QY 383 CACTAGTAGCGTGGTAGCTACTTGGCCCAACACTGTTATTGAGACCTTGCCTGTGAGTGG 442
Db 400 CACCAACAGCGTGGCAGCTACTCTGCCCAACACGCTGACCGACACTCGCGGGGAGCGG 459
QY 443 TGCATGATGACTAGTTGAGCGAGTGGCGACGACCTGTGCTGCTTACCTGCTGGCACA 502
Db 460 GCGTGGGGGCTGTGCTGCGCGGTGGGCGACAGCTGCTGTGTTCACTTGTGGCAG 519
QY 503 CTGTGCTCTTTATCTTCTGTGTCGCCCGCAGCTGTGCTTACCGAGGTGTGGGTCTCCCT 562
Db 520 CTGCGGNTTTTGTGCTGTGNTCCAGCTGCGCTTACCANGTGTGCGGCGCGCT 579
QY 563 GTACCAATTTGTCCACACGAGTATCTGGGCTCTGTGTCGGCTAGTTACAGGCCAC 622
Db 580 GTACAGCTCGGCGCTGCNACTCAGCGCGCGCCCGCCACACGCTANT----- 628
QY 623 CGAGCCGCTGGCGAGGAATTTCACTAACTTAGTTCTTACCAACAGATCAAGAGCAGTAG 682
Db 629 -----GGACCGAANGCTGTGGATCCAACGGCTTGGAACTAG 670
QY 683 TCGCAGGAGACACGAAACCCCTGGCTTGCCATCTCGAGGTACAAAGAGCATCTGAG 742
Db 671 -CGTCAGGAGCGCGGGCTCCCTCTGGGTGCGAGCCCGGGTGGCAGGAGCGCGGG 729
QY 743 TCTCACCAGTACAAGTGTGCTTCACTAAGAGGCGCAGATCTATCTGTCTCCGAGAGT 802
Db 730 CAGTCCAGCCGAGTCTGCGTTGTCACAGAGGCCCAGCGTGGCGCTGCCCCCTGAGCC 789
QY 803 GGAGGAGGGAACCCACAGGAGGTGTCTACAAACCCCATCAGGCAATCATGGGTGCCAAG 862
Db 790 GGAGCGGACGCGCGTGTGGCGAGGGGTCTGGGCGCCACCCCGGCGAGGACGCTTGGACCGAG 849
QY 863 TCCTGCTCGGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTTGTCTTCTTAA 913
Db 850 TGACCGTGGTTCTGTGTGTGTGTCACCTGCCAGACCGCGGAGAGGCCACTCTTTGGA 909
QY 914 AGGAAAGTGTCTGACCTGAGTCTCTCTGGGTGCGTGTCTGTAAACACAGACCCAGCTC 973
Db 910 GGGTGGCTCTCTGGCACGCGGCGCACTCCACCCCATCCGTGGCGCGCGACGCGCACGCGGG 969

```

974 CACATCTCTGCTGTCACACACCCGCCAAATGCTTTTCAGCTCAGGCCATTTATGAGAC 1033  
970 CCCCCATCCACATCGGGGCAACAGCTCTGGGACAGGCTTGTCCCGGTTGATGACC 1029  
1034 CAGACATTTCTTTACCTCAGGGGAGATGSCCAAGAGCGTCTAAACCCCTCATCTACT 1093  
1030 GAGACCAAGCACTTCTCTACTCTCAGGGGCAAAAGNACACTCGCNCCTCTCTCTACT 1089  
1094 CAGCA-ACCTCCAGCCTAACTTGACTGGGSCCAGGAGACTGGTGGAGATCACTTTCTGG 1152  
1090 CAATATATCTAGGCCCCAGCTGACTGGGCTTGGAGGTTCTGGAGACANTCTTTCTG 1149  
1153 GCTCAAGGCTTAGACATCAGGACCACTCTGCAAGGACACACCGCTCTATCGCTGATACT 1212  
1150 GTTCCAGGCTTGGATGCCAGGA-TTCCCGGCAAGTTGGCCGCTCTGCCACGCGTACT 1208  
1213 GGCAGATGCGGCCCTGTTCCACAGCTGCTGTGAACCATGACAGATGCCAATATGICA 1272  
1209 GGCAAATGCGGCCCTGTTTCTGAGCTGCTTGGGAACCAACGCGCAGTCCCTACGGGG 1268  
1273 GACTCTCAGGTCACTTGCAGGTTTCGAACAGCAAAACCAACAGGTGACAGATGCT--- 1329  
1269 TGTTCCTCAAGACGCACTGCCCGCTGGAGCTGGGTCAACCCAGCAGCGGTGCTGTG 1328  
1330 -----TGAAACCAAGCCCAACG 1346  
1329 CCGGGGAGAGCCCGGCTCTGTGGCGCCCGCCGAGGAGGAGGAAACACAGACCCCGT 1388  
1347 CACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTAATGCTTTCTCGG 1406  
1389 CGCTGTGTCAGTGTCTCCGACAGCAGCAGCCCTTGGCAGGTGTACGGTTCGTGGG 1448  
1407 GCCTGTCTTGCAAGGTGGTGTCTGTAGTCTCTGGGTTACCAAGCACAAATCAGCGCCG 1466  
1449 GCCTGCTGTGGCGGCTGTGGCGCCCGCCAGGCTCTGGGGTCCAGGCACAAACGACGCGC 1508  
1467 TTCTTTAAGAACTTAAAGAGTTTATCTCTGTGGGAATACGGCAAGTATCACTGAG 1526  
1509 TTCTCAGGAACACCAAGATTTATCTCCTGGGAAGATGCCAAGTCTCGCTGAG 1569  
1527 GAATGATGGAAGATGAAGTAGAGGATTCCTACTGCTCGCAGCAGCCCGGGAG 1586  
1569 GAGCTGACGTGAAGATGAGGTGGGAGCTGGCTTGGCTGGCGAGGAGCCAGGGTT 1628  
1587 GACGCTTCCCGCTGACAGACCCCTCTGAGGAGAGGATCTGCTACGTTCTCTGTT 1646  
1629 GCCTGTCTCCGGCCGACAGCACCGTCTGCTGAGGAGATCTGGCCAACTTCTGAC 1688  
1647 TGGCTGATGACACATACGTTGGTACAGCTGCTTAGGTCACTTTTATCATCAGAGAGC 1706  
1689 TGGCTGATGATGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTATGTCAGGAGAC 1748  
1707 ACATTCCAGAAAGACAGGCTCTTCTTACCTGPAAGTGTGTGAGGAGTGTGAGAGC 1766  
1749 ACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
1767 ATTGGAGTCAGGCAACACCTTGAGAGGTGCGGCTACGGGAGTGTCAACAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGACNTTGAAGAGGTGTGAGTCTCGGGAGCTGTGGAGCAGAGTC 1868  
1827 AGGCATCACCAGAGACCTGGCTAGGCATGCCCATCTGCAGACTCGGTTTATCCCAAG 1886  
1869 AGGCAGCATCGGGAAGCAGGCCCGCCCTGCTGACGCTCCAGACTCCGCTTCTATCCCAAG 1928  
1887 CCAAGCGGCTCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTGGGC 1946  
1929 CTTGACGGGCTGGGCCGATTTGTGAACATGAGTACGTCGTGGAGCAGAACGTTCCGC 1988  
1947 AGAAGGAGCAGGCCAGCAATTTACCCAGCGTCTCAGACTCTCTTTCAGCATGTCTAAC 2006  
1989 AGAGAAAGAGGCCGAGGCTCTCACCTCGAGGGTGAAGGCACCTGTTTCAGCGTGTCTAAC 2048  
2007 TATGAGCGGACAAAACATCTCACTTATGGGGTCTTCTGTACTGGGTATGATGACATC 2066

2049 TAGAGCGGCGCGCGCCCGGCTTCTGGGCGCTCTGTCTCGGCTGACGATATC 2108  
2067 TACAGACCTTGGCGGCCCTTTGTGTGCTGTGCTGTGCTGTGACACACACCCAGGATG 2126  
2109 CACAGGCGCTGGCGCACCTTCTGTCTGTGCTGTGCGGGCCAGGACCCGCGCTGAGCTG 2168  
2127 TACTTTTAAAGCAGATGACCGCGGCTATGATGCATCCCGCCAGGTTAAGTGTGTG 2186  
2169 TACTTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCCAGGACAGGCTCAG 2228  
2187 GAGGTTTGTCCCAATATGATCAGGCACCTCGGAGACAGTACTGTATCCGCCAGTATGCA 2246  
2229 GAGTCATGSCCAGCATCATCA---AACCCAGAACAGTACTGCTGCTGCTATGCC 2285  
2247 GTGGTCCGAGAGATAGCAAGGCCAAATGCCACAGTCTTTTAGAGACAGGTACACACC 2306  
2286 GTGGTCCAGAAAGGCCCCCATGGGCAGCTCCGCAAGGCTTCAAGAGCCACGTCTCTTACC 2345  
2307 CTCTCTGACTCCAGCCATACATGGGCCAGTTCTTAAAGCATCTGCAGATTCAGATGCC 2366  
2346 TTGACAGACTCCAGCGGTACATGCGACAGTTCTGTGGCTCACCTGCAG-----GANAAC 2399  
2367 AGTGCACTGAGGAACCTCGTTGTCTATGAGCAGACATCTCTATGAATGAGAGCAGCAGC 2426  
2400 AGCCCGCTGAGGATGCGCTGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAGCAGT 2459  
2427 AGCTGTTTGTACTTCTCTGACATCTCTGCTGACAGTCTGCTGCTGCTGCTGACAGG 2486  
2460 GGCTCTTTCAGCGTCTTCTACGCTTCAITGTGCCACAGCGCTGCGCATCAGGGGCAAG 2519  
2487 TGCTATACGCACTGCGCAGGCAATCCCGCAGGCTCCAGCCTATCCACCTCTCTCTGAGT 2546  
2520 TCTAGCTCCAGTGGCGAGGATCCCGCAGGCTCCATCTCTCCACGCTGCTCTGACG 2579  
2547 CTGTGTTTCGGAGACATGAGAACAAAGCTGTTGCTGAGGTGCAGCGGAGTGGTGGT 2606  
2580 CTGTGCTACGGGCAATGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGGCTGCTC 2639  
2607 TTACGTTTGTGATGACTTCTGTGTGAGCGCTCACTTTGGACCAAGCAAAACCTTC 2666  
2640 CTGCGTTTGTGATGATTTCTTGTGTGAGCACTCACTCACCCACGCGAAACCTTC 2699  
2667 CTCAGCACCTCGTCCATGGCGTCTCTGAGTATGGTGATGATTAAGTTCAGAGACA 2726  
2700 CTCAGACCTCGTCCGAGGTGCTCCTGAGTATGGCTGCTGCTGCTGCTGCTGCTGAGCA 2759  
2727 GTGGTAACTTCCCTGTGAGCGCTGCTACCTCGGCTGGTGAGCTCCATACACAGCTGCT 2786  
2760 GTGGTAACTTCCCTGTAGAGACAGAGCGCTTGGTGACAGGCTTTTGTTCAGATGCCG 2819  
2787 GCTCACTGCTGTTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2846  
2820 GCCACGGCTTATCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2879  
2847 TGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCGCTCACTTCCAGAGTGC 2906  
2880 AGGCACTACTCAGCTATGCCCGGACCTTCAATCAGAGCCAGTCTCACTTCAACCGGCG 2939  
2907 TTTCAAAGCTGGGAAGCAATGCGGAACAAAGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCT 2966  
2940 TTTCAAGCTGGGAGAAACATGCTGCGCAAACTCTTTGGGGTCTTTGGGCTGAAAGTGCAC 2999  
2967 GGTCTATTTCTAGACTTCCAGGTGAGACCGCTCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3026  
3000 AGCCTGTTTCTGATTTGAGGTGAGACCGCTCCAGAGCGGTGCTGCAACCAACATATCAAG 3059  
3027 ATCTTCTCTGCTTCCAGGCTTACAGGTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3086  
3060 ATCCCTCTGCTGACGCGTACAGGTTTTCACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3119  
3087 CGTGTAGGAAGAACTCAGATCTTCTGCGGCAATCTCTCCAGCCAGGATCTCTGCTGCTGCTGCTGCT 3146



803 GGAGGGGACCCACAGGAGGTGTACCAACCCCATCAGGCAATCATGGTCCCAAG 862  
790 GGAGCGAGCCCGTTGGGAGGGGCTCTGGGCCCAACCCGGGAGGACGCCCTGGAGCCGAG 849  
863 TCTGTCTCGGTCCCCCGAGTGCTACT-----GAGAGAGAAGATTCTCTTTCTAA 913  
850 TGACCGTGGTTTCTGTGTGTGTCACTGCGACAGCCGCGGAGAGCCACCTCTTTGGA 909  
914 AGGAAAGGTCTGTGACCTGAGTCTCTCTGGTGGTGTGCTGTAAACACAAAGCCAGTCTC 973  
910 GGGTGCCTCTCTGGGACCGGCACCTCCCAACCATCCGTGGGCGCGCAGCACACCGGG 969  
974 CACATCTCTGTGTACCAACCCCGCCAAAATGCTTTCAGCTCAGGCCAATTTATTAGAC 1033  
970 CCCCCCATCATCTCGGGCCACACAGTCTGGGACACGCTTGTCCCCCGGTGTACGCC 1029  
1034 CAGACATTTCTTTACTTCAGGGGAGATGGCCAAAGAGGTCTTAAACCCCTCATTTCTACT 1093  
1030 GAGACCAAGCACCTTCTCTACTCTCTCAGGCGCAAGNACACTGCGNCCCTCTCTTCTACT 1089  
1094 CAGCA-ACCTCCAGCCTAACTTGACTGGGGCCAGGAGACTGGTGGAGATCATCTTTCTGG 1152  
1090 CAATATATCTGAGGCCACAGCTGACTGGCGTTGGGAGGTTCGTGGAGACATCTTTCTG 1149  
1153 GCTCAAGCCTAGACATCAAGACCACTCTGCAAGGACACACCGTCTATCGGTGCGATACT 1212  
1150 GTTCCAGGCTTGGATGCCAGGA-TTCCCCGCAAGTTGGCCCGCTTGCACCAGCGTACT 1208  
1213 GGCAGATGGGCCCCCTTCCACAGCTGCTGTGAACCATGCAAGAGTCCCAATATGTCA 1272  
1209 GGCANAATGCGCCCCCTGTTCTGAGAGTCTTGGGAACCAAGCGGAGTCCCGCTACGGG 1268  
1273 GACTCCTCAGGTCAATTTGAGTTCGAAACAGCAAAACCAAGGTTGACAGATGCGCT--- 1329  
1269 TGTTCCTCAAGACGCACTGCGCGCTGGAGCTGCGGTCAACCCAGCAGCGGTGTCTGTG 1328  
1330 -----TGAAACACAGCCCAACCG 1346  
1329 CCGGGAGAACCCCGGCTGTGTGGGCGCCCCGAGGAGGAGGAACACAGACCCCGT 1388  
1347 CACCTCATGATTTGCTCCGCTGCAAGCAGGTCCCTGGCAGGTATATGGTTTCTTGG 1406  
1389 GCGCTGTGAGCTGCTCCGCCAGCAGCAGACCCCTTGGCAGGTGTACGGTCTGTGG 1448  
1407 GCCTGTCTGCAAGGTGGTGTCTGTAGTCTCTGGGTACAGGACAAATGAGCGCGC 1466  
1449 GCCTGCTGCGCGCGCTGGTGCOCACAGGCTCTGGGGCTCCAGGCAACAGAACCGCGC 1508  
1467 TTCTTTAAGAACTTTAAAGAGTTTCATCTCTGGGAAATACGGCAAGCTATCACTGCAG 1526  
1509 TTCTTCAGGAACCAAGAGTTTCATCTCTGGGNAAGCATGCCAAGTCTGTGTGAG 1568  
1527 GAATGATGTGAAGATGAAGTAGAGATTGCCATGGCTCCGCAAGCAGCCCGGGAG 1586  
1569 GAGCTGACGTGGAAGATGAGGTGCGGAGTGGCTGGCTGGCGAGGAGCCAGGGGT 1628  
1587 GACCGTGTCCCGCTGAGAGCACCGTCTCAGGAGAGGATCTGTGCTAGTCTCTGTTC 1646  
1629 GCTGTGTTCGGCGCAGAGACCCGTCTGCTGTGAGATCTGTGCGCAAGTCTCTGCAC 1688  
1647 TGGCTGATGGACACATACGTGGTACAGTCTTAGTCTATTTTATCATCAGAGAGC 1706  
1689 TGGCTGATGAGTGTACGTCTGAGCTGCTCAGGTCTTTTCTTTTATGTCAGGAGACC 1748  
1707 ACATTCAGAGACAGGCTCTTCTTCTACCTAGGTGTGTGGAGCAAGCTGCGAGC 1766  
1749 ACGTTTCAAAAGACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
1767 ATTGAGTTCAGGCAACACCTTTGAGAGAGTTCGGGTACGGAGGTGTCAAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGCACTTGAAGAGGAGTTCAGTTCGGGAGCTGTTCGGAAGCAGAGTTC 1868

QY 1827 AGGCATCACGAGACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCCAAG 1886  
Db 1869 AGGCACATTCGGAAAGCCAGGCCCGCTCTGTGAGCTCCAGACTCCGCTTCATCCCAAG 1928  
QY 1887 CCACAGGCTCGCGGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTTGGGC 1946  
Db 1929 CTTGACGGGCTCGCGCGATTGTGAACATGAGTACGTCTGTGGAGCCAGAACCTTCGC 1988  
QY 1947 AGAAGAAAGCAGGCGCCAGCATTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCTCAAC 2006  
Db 1989 AGAAGAAAGGCGCCAGCGTCTACCTCGAGGGTGAAGGCACCTGTTCAGCGTGTCTAAC 2048  
QY 2007 TATGACGCGACAAAACATCTCTATGCGGTCTTCTGTACTGGGTATGAATGACATC 2066  
Db 2049 TACGAGCGGCGCGCGGCCCGGCTCTCTGGCGGCTCTGTGTGGGCTTGAGCATATC 2108  
QY 2067 TACAGACCTGGGGGCGCTTGTGTGCTGTGGTGTCTCTGGACACAGACCCAGATG 2126  
Db 2109 CACAGGGCTGGCGCACCTTGTGTGTGTGTGGGCGCCAGGACCCCGCCCTGAGCTG 2168  
QY 2127 TACTTTGTTAAGCGAGATGTGACCGGGGCTATGATGCCATCCCCCAGGGTAAGCTGTG 2186  
Db 2169 TACTTTGTCAAGGTGAGTGTGACGGGCGGTACGACACCATCCCCCAGGACAGGCTCAG 2228  
QY 2187 GAGGTGTTGCGAATATGATCAGGCACCTCGGAGAGCAGTACTGTATCCGCCAGTATGCA 2246  
Db 2229 GAGGTTCATCGCCAGCATCATCA---AACCCAGAAACAGTACTGCTGCTGCTGCTATGCC 2285  
QY 2247 GTGGTCCGAGAGATAGCCAGGCCAAGTCCACAGTCTTTAGGAGACAGTCAACACC 2306  
Db 2286 GTGGTCCAGAGGCCGCCCATGGGCACTGCGCAGGCTTCAGAGCCACGTCTCTACC 2345  
QY 2307 CTCTGTGACTTCAGCCATACATGSGCCAGTTTCCTTAAGCATCTGACAGATTTCAGATGCC 2366  
Db 2346 TTGACAGACTTCAGCCGTACATGCGACAGTTTCGTGGTCTCACTGCAG-----GANAAC 2399  
QY 2367 AGTGCACTGAGGAACCTCGTGTGTCATCGAGCAGAGCATCTCTATGAATGAGAGCAGCAGC 2426  
Db 2400 AGCCCTCAGGAGTGCCCTGTCTATCGAGCAGAGCTCTCTCCTGAATGAGGCCAGCAT 2459  
QY 2427 AGCCTGTTGACTTCTCTGCACTTCTCTGCGTCAAGTGTCTGAGGTGCAAGGATGGTGACAG 2486  
Db 2460 GGCCTCTTGGACGTCTTCTACGCTTCATGTGCCACACCGCGCTGCGCATCAGGGGCAAG 2519  
QY 2487 TGCTATACGAGTGCAGGCGCATCCCGCAGGCTCCAGCCTATCCACCCCTGCTGTGAGT 2546  
Db 2520 TCTACGTCCAGTGCAGGGGATCCCGCAGGCTCCATCTCTCTCAAGCTGCTCTGACG 2579  
QY 2547 CTGTGTTCCGAGACATGAGAAACAAGCTGTTTGTGAGGTGCAAGCGGATGGGTGGTTC 2606  
Db 2580 CTGTGCTACGGCGACATGGAGAAACAAGCTGTTTGGGGGATTTCGGCGGACGGCTGCTC 2639  
QY 2607 TTAGGTTTGTGATGACCTTCTGTGTGAGCGCTCTCTGGACCAAGCAAAACCTTC 2666  
Db 2640 CTGCGTTTGTGTGATTTCTTGTGTGACACCTCACCTCACCCAGCGCAAAACCTTC 2699  
QY 2667 CTGAGCACCTGTCTCATGCGCTTCTGAGTATGGTGCATGATATACTTCAGAGAACA 2726  
Db 2700 CTGAGACCTGCTCGAGGTGCTCTGAGTATGCTGCTGGTGGTGAACCTTGGGAGAC 2759  
QY 2727 GTGGTGAACCTTCCCTGTGAGCCTGTGACCTCGGTGTGTGAGCTCCATACAGTGCCT 2786  
Db 2760 GTGGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGGGCTTTTGTTCAGATGCC 2819  
QY 2787 GCTCACTGCTGTTTCCCTGTGTGCTTGTCTGCTGGACACTCAGACTTTTGGAGTGTTC 2846  
Db 2820 GCCCAGGCGCTATTTCCCTGTGTGCGGCTCTGTCTGTGATACCCGAGCCCTGGAGTGCAG 2879  
QY 2847 TGTGACTACTCAGGTATGCCCCAGACCTCAATTAAGACGAGCCTCACCTTCAGAGTGTTC 2906  
Db 2880 AGCGACTACTTCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGC 2939  
QY 2907 TTCAGACTGGGAGAGACCATGCGGAAACAGCTCTCTGTGCTTGTGGGTTCAGTGTTCAC 2966



2940 TTCAAGCTGGAGAGCAATCGCTCGCAACTCTTTGGGCTCTGGGCTGAAGTGTAC 2999  
2967 GGTCTATTCTAGACTTCAGAGTGAAGAGCTCCAGACAGTCTGACATCAATATACAAG 3026  
3000 AGCTGTCTTGGATTTCAGGTGAAGAGCTCCAGACGGTGTGACCAACATCTACAAG 3059  
3027 ATCTTCCTGCTTCAGGCTACAGTTCATCATGTGTGATTCAGTTCCTTTGACCAAG 3086  
3060 ATCTTCCTGCTTCAGGCTACAGTTCATCATGTGTGATTCAGTTCCTTTGACCAAG 3119  
3087 CGTGTAGGAAGAACTCACATTTCTTTGCGCATCATCTCCAGCCAAAGCATCTCTGTC 3146  
3120 CAAAGTTTGAAGAGACCCACATTTCTCTGCGCTCATCTCTGACAGGGCTCCCTCTGC 3179  
3147 TATGTATCTTGAAGTCAAGATTCAGGATGACACTAAG-----GCTCTGGC 3197  
3180 TACTCCATCTTAAAGCCAAAGACGACGGATGTGCTGGGGGCCAAGGGCGCGCGGC 3239  
3198 TCCTTTCTCTGAGAGCGCACATTTGGCTCTGTACCAAGGCTTTCTGCTCAAGCTGGCT 3257  
3240 CTCTGCTCTCGAGCGCTGACGTGGTGTGCGCACCAAGCATTTCTGCTCAAGCTGACT 3299  
3258 GGTATCTGTCATCTCAATATCTCTCTGGACCTCTGAGGACAGCCCAAAATCTGCTG 3317  
3300 CGACACCGGTGTCACTAGTGTGCTCTCTGGGTCACTCTGAGGACAGCCCAAGCGAGCTG 3359  
3318 TGCCGAGAGCTCCAGAGCGCACATGACCATCTTAAAGCTGAGCTGACCCAGCCCTA 3377  
3360 ATGCGAAGCTCCCGGAGCAGCGCTGACTGCTGAGGCGCGGCGGACCAACCGGACTG 3419  
3378 AGCAGAGCTTTTCAAGACATTTTGGACTPA 3407  
3420 CCTCAGACTTCAAGACCATCTGACTGA 3449

## BULT 15

-09-430-323-173  
Sequence 173, Application US/09430323  
Patent No. 6309867

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.

TITLE OF INVENTION: No. 6309867el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/430,323

FILING DATE: 29-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 173:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4029 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: 1..4029  
LOCATION: /note= "preliminary sequence for  
human TET cDNA insert of  
Plasmid pORN121"  
SEQUENCE DESCRIPTION: SEQ ID NO: 173:  
US-09-430-323-173

Query Match 43.5%; Score 1520.2; DB 4; Length 4029;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;

QY	23	CCCGGCTT	GAGCACA	TAATGAC	CCCGCGCT	CTCTGTT	CCCCCGCG	GTGCGT	CTCTGTT	GGG	82
DB	40	CCCGGCCA	CCCCCGCA	TGCGCG	CGCTCC	CGCTCC	CGAGCG	GTGCGT	CTCTGTT	GGG	99
QY	83	CAGCGGAT	ACCGGAG	GTGGCG	CGCTTGG	CAACCTT	TGTGCG	CGCGCT	TGGGCG	CGCGAG	142
DB	100	CAGCCACT	TACCGAG	GTGTGCG	GTGGCC	AGCTTGT	GTGGCG	CGCTTGGG	CGCGCG	CGCGAG	159
QY	143	CAGCGGCT	TGTCAAC	CCCGGGA	CCCGAAG	ATCTACC	GCACCTT	TGGTGG	CGCAAT	GCCT	202
DB	160	CTGGCGGT	GTGTGAG	CGCGGGA	CCCGCG	GGCTTTC	CGCGCG	GTGTTGG	TGGCGCG	CGT	219
QY	203	AGTGTGAT	GCATG	TGGGGCT	CACAGCT	CCACCT	CGGACCTT	TCCTTCC	TCCAC	CGAGTGC	262
DB	220	GGTGTGCT	GTGCTT	GAGCTT	GTACAG	AGGCG	CAGAGG	CGGCGCT	CCCATG	CGCTT	279
QY	263	ATCCCTGA	AGAGCT	GTGGCC	CAGGGT	TGTGCA	GAGACT	CTGCG	AGCGCA	ACGAGAGAA	322
DB	280	CTGCTGA	ANGANT	CTGGTG	CGCCAG	TGTGCT	GCANANG	TGTGCG	ANGC	GTGCGGCGCGGCA	339
QY	323	CGTGTGCT	TTTGGCT	TTGAGCT	GCTTAC	AGAGG	CCAGAGG	CGGCGCT	CCCATG	CGCTT	382
DB	340	CGTGTGCT	TTGCGCT	TTGCGCT	TTGCGCT	TTGCGCT	TTGCGCT	TTGCGCT	TTGCGCT	TTGCGCT	399
QY	383	CACGTAGT	AGCTGT	GTAGTACT	TGCCCC	AACTGT	TATTTAG	ACCCCT	CGTGT	CGAGTGG	442
DB	400	CACACCA	CGCTGT	GCAGCT	ACTCTG	CCCCAC	ACACG	GTGAC	CCGCA	CTGCGGGGAGCGG	459
QY	443	TGCATGAT	GTCTACT	GTGTGAG	CGGAG	GTGGCG	ACGACCT	GTGTGCT	GTGTCT	ACTGTTG	502
DB	460	GCGGTGGG	GTGTGCT	GTGTGCT	GTGTGCT	GTGTGCT	GTGTGCT	GTGTGCT	GTGTGCT	GTGTGCT	519
QY	503	CTGTGCT	TTTATCT	TTTCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	562
DB	520	CTGCGCG	NTNTT	TTGCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	TTGCTG	579
QY	563	GTACCAAA	TTTGTG	CCACCG	AGGATAT	CTGCG	CCCTCT	GTGTG	CTGCTAG	TTTACAGG	622
DB	580	GTACCACT	GTGGCG	CTGCNA	CTCAGG	CGCGCG	CGCGCG	CGCGCG	CGCGCG	CGCGCG	628
QY	623	CGGACCC	CGTGGG	CAGGAAT	TTTCACT	TAACCT	TTAGT	TTTAC	ACAGAT	CAAGAG	682

629 -----GGACCCGAAGCGTCTGGATCCACGGGCTTGGACCATAG 670  
683 TGCACGGAAGCACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 742  
671 -CCTCAGGAGACCGGGTCCCTTGGCTGCCAGCCCGGGTGCAGAGAGCGCGGGG 729  
743 TCTCACCACTACAGTGTGCTTCAAGTAAAGAGCCAGATGCTATCTGTCCCGAGAGT 802  
730 CAGTGCACCGCAAGTCTGCGCTGCCCAAGAGGCCAGGCGTGGCGTGCCTGAGCC 789  
803 GAGAGAGGACCCACAGCAGCGGTGTACCAACCCCATCAGCGCAATCATGGGTGCCAAG 862  
790 GAGAGCGACCGCGTGTGGCAGGGGTCTGGGCCCCACCCGGGCGAGCGCTGACCCGAG 849  
863 TCCTGTCTGGTCCCGGAGTGCCTACT-----GCAGAGAAAGATTGTCTTTAA 913  
850 TGACCGTGTCTGTGTGGTGTACCTTGCACAGCCCGCGAGAGGCCACTCTTTGGA 909  
914 AGGAAAGGTGTGACCTGAGTCTCTGTGGGTGGGTGTGTGTAAACAAGCCGACTC 973  
910 GGTGTGGCTCTGTGGCAGCGCACTCCACCCATCCGTGGGCGCCAGCACACCGGGG 969  
974 CACATCTGTGTGCACACCGCGCAAAATGCTTTTGTGCTCAGCTCAGGCCATTATTGAGC 1033  
970 CCCCCCATCCATCCGCGCCACACGCTCTGGGACAGCGCTTGTGCCCGGTGTACGCC 1029  
1034 CAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGGGTCTTAAACCCCTCATTTCTACT 1093  
1030 GAGACCAAGCACTTCTCTACTCTCCTCAGCGGACAAAGNACACTTGCNGNCCCTCTTCTACT 1089  
1094 CAGCA-ACCTCAGCCTAACTTGACTGGGGCCAGGAGACTGGTGAGATCATCTTTCTGG 1152  
1090 CAATATATCTGAGGCCACGCTGACTGGGCTTTCGGGAGGTTTGTGGAGACANTCTTCTG 1149  
1153 GCTCAAGGCTAGGACATCAGACCACTCTGAGGACACACCGCTCTATCGCGTGTGACT 1212  
1150 GTTCCAGGCTTGGATGCCAGGA-TTCCCGGAGGTTCGCCCGCTGCCCGAGTACT 1208  
1213 GGAGATGGGCCCCCTGTTCGAAGCTCTGGTGAACCAATGAGAGTGCATATGTCA 1272  
1209 GGCAAAATGGGCCCCCTGTTCGAGCTCTTGGGAAACCAAGCGAGTGGCCCTACGGGG 1268  
1273 GACTCTCAGTGCACATTCAGGTTTCGAACAGCAAAACCAAGAGTGCATGCCT--- 1329  
1269 TGTTCCTCAAGAGCACTCCCGCTCGAGCTTGGGTACCCAGAGCCGCTGTCTGTG 1328  
1330 -----TGAAACACAGCCCAACG 1346  
1329 CCGGGAGAGACCCAGGCTCTGTGGGCGCCCCGAGGAGGAGAAACACAGACCCCGT 1388  
1347 CACTCATGATTTGCTCGCTGCACAGAGTCCCTGGCAGGTATATGTTTCTTCGG 1406  
1389 CGCTGGTGAGTGTCTCGCCAGACAGCAGCCCTGGCAGGTGTACGGCTTCGTGCGG 1448  
1407 GCCTGTCTCTGAAGTGTGTCTGTAGTCTCTGGGTACAGGCAAAATCAGCGCCG 1466  
1449 GCCTGCTGCGCGGTGTGTGCCCCAGGCTCTGGGCTCCAGGCAACAGAGCGCG 1508  
1467 TCTTTAAGAACTTAAGAGTTTCTGTGTGGGAAATACGGCAAGCTATCATCTGAG 1526  
1509 TTCTCAGGAACACCAAGAAGTTTCTCTCCCTGGGGAAGCATGCCAAGCTCTCGTGCAG 1568  
1527 GAATGATGTGAAGATGAAGTGAAGGATTCGCACTGGCTCCGAGAGCCCGGGGAG 1586  
1569 GAGTGAAGTGAAGATGAGCTGTGGGAGCTGCGCTTGGCTCCGAGAGCCAGGGTT 1628  
1587 GACCGTGTCCCGCTCAGAGCAACCGTCTGAGGAGAGATCCTGGCTACGTTCTGTTC 1646  
1629 GGCTGTGTTCCGGCCGAGAGACCGCTCTGGTGAGGAGATCCTGGCCAAAGTCTCTGCAC 1688  
1647 TGGCTGATGACACATACGTGTGACGCTGTAGGTCAATCTTTTATCATCAGAGAGC 1706  
1689 TGGCTGATGATGTGACGTGTCGAGCTGCTCAGTCTTCTTCTTTATTGTACGGAGACC 1748

1707 ACATTCAGAGAACAGCGCTTCTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCAGAGC 1766  
1749 AGTTCCTAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
1767 ATTGGAGTCAGGCAACACCTTGAAGAGTGGGCTACGGAGTGTGTCAAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGCACTTGAAGAGGTGAGTGGGAGCTGTCCGAAGCAGAGGTC 1868  
1827 AGGATCACACAGGACACTGCTAGCCATCCCATCTGACAGCTGCGCTTCTCCCAAG 1886  
1869 AGGACAGATCCGGAGAGCAGGCCCGCTCTGACGTCCAGATCTCCGTTTCACTCCCAAG 1928  
1887 CCCAACGSCCTCGCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACCAAGAGTTTGGG 1946  
1929 CTTGAGCGGCTCGCGCCGATTGTGAACATGGAATCTGCTGGGAGCCAGAACGTTCCCG 1988  
1947 AGAGGAGAGAGGCGCCAGCATTTACCCAGGCTCTCAAGACTCTCTTCAGCATGCTCAAC 2006  
1989 AGAGAAAGAGGGCGGAGCGCTCACCTCAGGGGTGAAGGCACTGTTTACGCTGCTCAAC 2048  
2007 TATGAGCGGCAAAAACATCTCACTTATGGGGTCTTCTGTACTTGGGTATGAATGATC 2066  
2049 TACGAGCGGGCGCGGCCCGGCTCTCTGGGCGCTCTGTGTGGGCTTGGAGATATC 2108  
2067 TACAGGACCTGGCGGCGCTTTGTGCTGCTGTGGTGTCTCTGACACAGACCCAGGATG 2126  
2109 CACAGGGCTGGCGCACCTTCTGCTGTGTGGGCGCCAGGACCGCGCGCTGAGCTG 2168  
2127 TACTTTTAAAGGAGATGTGACCGGGCGCTTATGATGCCATCCCGAGGTTAAGCTGGTG 2186  
2169 TACTTTGTCAAGTGTGATGTGACCGGGCGGTAACACACATCCCGCAGGACAGCTCACG 2228  
2187 GAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATTCGCCAGTATGCA 2246  
2229 GAGGTTCATCGCCAGCATCATCA---AACCACAGAACACGTACTGCTGCTGCTGATGCC 2285  
2247 GTGGTCCGAGAGATAGCCAGGCGCAAGTCCACAGTCCCTTAGGAGACAGGTCACACC 2306  
2286 GTGGTCCAGAAAGCGCCCATGGGCAAGTCCGAAAGGCTTCAAGAGCCACGCTCTTACC 2345  
2307 CTCTCTGACCTCCAGCCATACATGGGCGCATCTTAAAGCATCTGACAGGATTCAGATGCC 2366  
2346 TTGACAGACCTCCAGCGTACATCGACAGTTTGGTCTCACCTGCAG-----GANAAC 2399  
2367 ATGCACTGAGAACTCCGTTGTATGAGCAGAGCATCTATGATGAGAGCAGCAGC 2426  
2400 AGCCGCTGAGGAGTGGCGTGTCTATGAGCAGAGCTCCTCCCTGATGAGGCGCAGT 2459  
2427 AGCCTGTTTGACTTCTTCTGCACTTCTGCTGTCACAGTGTCTGTAAGATTGGTGACAG 2486  
2460 GGCCTCTTCGAGCTTCTTACGCTTCTATGTGCCACACCGCGTGGCCATCAGGGGCAAG 2519  
2487 TGTATACGCAAGTCCAGGCGCATCCCGCAGGCTTCCAGCTTATCCACCTGCTCTGAGT 2546  
2520 TCTTACGTCAGTGCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTGAGC 2579  
2547 CTGTCTTTCGGGACATGGAGAACAGCTGTTGCTGAGTGCAGCGGAGTGGTTCCTT 2606  
2580 CTGTCTACGGCGCATGGAGAACAGCTGTTTCGGGAGTTCGGCGGAGCGGCTGCTC 2639  
2607 TTACCTTTTGTGATGACTTCTGTGTGACGCTCTTGGACCAAGCAAAACCTTC 2666  
2640 CTGCGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCACCCACGCGAAACCTTC 2699  
2667 CTCAGACCTGCTGCTTCCATGGGTTCTCTGAGTATGGGTGATGATTAACCTTGCAGAGACA 2726  
2700 CTCAGAGCCCTGGTCCGAGGTCTCTCTGAGTATGGCTGCTGTGTGAACCTTTCGGAGAGACA 2759  
2727 GTGGTGAACCTTCCCTGTGGAGCTTGGTACCTTGGTGTGAGCTTCCATACCAAGCTGCTC 2786  
2760 GTGGTGAACCTTCCCTGTGAGAGAGCAGGCGCTTGGTGGCAGCGCTTTTGTTCAGATGCCG 2819



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: February 2, 2004, 01:51:02 ; Search time 1398 seconds  
(without alignments)  
9113.998 Million cell updates/sec

le: US-09-042-460-1  
fect score: 3496  
uence: 1 GAATTCCGGTGGAGGCC.....CCGAGTCGTGACCAAGCTT 3496

ring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

ched: 2434939 seqs, 1822278265 residues  
al number of hits satisfying chosen parameters: 4869878

imum DB seq length: 0  
imum DB seq length: 2000000000

t-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1600.8	45.8	4015	9	US-09-733-294A-3
2	1600.8	45.8	4015	10	US-09-990-080-1
3	1600.8	45.8	4015	10	US-09-843-676-224
4	1600.8	45.8	4015	10	US-09-953-052-1
5	1600.8	45.8	4015	15	US-10-053-758-224
6	1600.8	45.8	4015	15	US-10-208-243-1
7	1600.8	45.8	4015	15	US-10-054-295-224
8	1600.8	45.8	4015	15	US-10-054-611-224
9	1600.8	45.8	4015	15	US-10-105-963-1
10	1600.8	45.8	4015	15	US-10-044-539-1
11	1600.8	45.8	4015	15	US-10-044-539-1
12	1600.8	45.8	4027	12	US-10-385-882-1
13	1594.6	45.6	3396	10	US-09-749-728B-32
14	1593.6	45.6	3453	15	US-10-205-629-1
15	1593.6	45.6	13766	13	US-10-105-616-1

c	16	1584.8	45.3	8742	13	US-10-105-616-6	Sequence 6, Appli
	17	1520.2	43.5	4029	10	US-09-843-676-173	Sequence 173, App
	18	1520.2	43.5	4029	11	US-09-438-486-173	Sequence 173, App
	19	1520.2	43.5	4029	15	US-10-053-758-173	Sequence 173, App
	20	1520.2	43.5	4029	15	US-10-054-295-173	Sequence 173, App
	21	1520.2	43.5	4029	15	US-10-054-611-173	Sequence 173, App
	22	1323.2	37.8	3855	15	US-10-044-692-4	Sequence 4, Appli
	23	1323.2	37.8	3855	15	US-10-044-539-4	Sequence 4, Appli
	24	1022	29.2	1856	15	US-10-294-778-11	Sequence 11, Appli
	25	724.4	20.7	1314	15	US-10-294-778-9	Sequence 9, Appli
	26	716	20.5	2176	15	US-10-044-692-3	Sequence 3, Appli
	27	716	20.5	2176	15	US-10-044-539-3	Sequence 3, Appli
	28	710	20.3	1311	15	US-10-294-778-1	Sequence 1, Appli
	29	689.8	19.7	2171	10	US-09-843-676-100	Sequence 100, App
	30	689.8	19.7	2171	10	US-09-766-253-100	Sequence 100, App
	31	689.8	19.7	2171	11	US-09-438-486-100	Sequence 100, App
	32	689.8	19.7	2171	15	US-10-053-758-100	Sequence 100, App
	33	689.8	19.7	2171	15	US-10-054-295-100	Sequence 100, App
	34	689.8	19.7	2171	15	US-10-054-611-100	Sequence 100, App
	35	470.2	13.4	873	13	US-10-282-960-1	Sequence 1, Appli
	36	468.8	13.4	4200	15	US-10-044-692-6	Sequence 6, Appli
	37	468.8	13.4	4200	15	US-10-044-539-6	Sequence 6, Appli
	38	468.8	13.4	15418	9	US-09-783-203-1	Sequence 1, Appli
	39	468.8	13.4	15418	10	US-09-994-427A-1	Sequence 1, Appli
	40	468.8	13.4	15418	11	US-09-995-419A-1	Sequence 1, Appli
	41	468.8	13.4	15418	15	US-10-141-220-1	Sequence 1, Appli
	42	468.8	13.4	15418	15	US-10-023-969-1	Sequence 1, Appli
	43	468.8	13.4	15418	15	US-10-206-447-1	Sequence 1, Appli
	44	468.8	13.4	51552	9	US-09-733-294A-30	Sequence 30, Appli
	45	338.2	9.7	519	13	US-10-282-960-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-09-733-294A-3  
; Sequence 3, Application US/09733294A  
; Patent No. US20020045588A1  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: William Gaarde  
; APPLICANT: Susan M. Freier  
; APPLICANT: Edward V. Wanciewicz  
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
; FILE REFERENCE: ISPH-0527  
; CURRENT APPLICATION NUMBER: US/09/733,294A  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: 09/572,423  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 108  
; SEQ ID NO 3  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)...(3454)  
US-09-733-294A-3

Query Match	45.8%	Score	1600.8	DB	9	Length	4015
Best Local Similarity	68.9%	Pred.	No. 0				
Matches	2379	Conservative	0	Mismatches	967	Indels	108
Gaps	9						
Qy	23	CCCGGCGTTGAGCACAAATGACCGCGCTCTCGTGGCCCGCGGTGCGCTCTCTCTCGG	82				
Db	40	CCCGGCGCACCGCGCGATGCGCGGCTCCCGGTCGCGAGCGCGTCCCTCTCTCGG	99				
Qy	83	CAGCGATACCGGAGGTGGCGGTGCACACTTTGTGCGCGCCTGGGCGCCGAGGG	142				
Db	100	CAGCCACTACCGGAGGTGCTGCGGTGGCCACGTTCTGTGCGCGCCTGGGCGCCGAGGG	159				
Qy	143	CAGCGCGCTTGTGCAACCGCGGGAGCCGGAAGATACCGCACTTGTGTGTCGCCAATGCCT	202				

160 CTGGCGCTGCTGAGCGGGGACCGGCGCTTTCCGCGCTGCTGAGCCAGTGCCT 219  
203 AGTGTGATGCACTGGGGCTCAGACGCTGCACCTGCGCACTTTCCTTCCACAGGTGTC 262  
220 GGTGTGCTGCTGGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279  
263 ATCCCTGAAGAGCTGCTGGCGGAGGCTGTGTCAGAGACTCTGCGAGGCGCAACAGAGAAA 322  
280 CTGCTGAGAGAGCTGTGGCCCGAGTGTCTGAGAGGCTGTGCGAGCGCGCGCGGAGAA 339  
323 CGTGTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382  
340 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
383 CACTAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442  
400 CACCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
443 TGCAATGATGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502  
460 GCGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519  
503 CTGTGCTCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562  
520 CTGCGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579  
563 GTACCAAAATTTGTCACCAACGAGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 622  
580 GTACCAAGCTCGGCGCTGCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639  
623 CGAACCCGCTGGGAGGAATTTCACTAACTTAGGTTCTTAAACAGATCAAGAGCACTAG 682  
640 GCGTCTGGGATGC-----GACCGGCGCTGGAAACCATAG 672  
683 TCGCAGGAGAGACCGAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742  
673 CGTACAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732  
743 TCTACAGTACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802  
733 CAGTGCAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792  
803 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 862  
793 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 852  
863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
914 AGGAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970  
913 GGGTGGCTCTCTGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972  
971 CTCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCCCCATCAGATCGGGGCCACAGTCTCTGCGACAGCGCTTCTGCTGCTGCTGCTGCTGCT 1032  
1028 TGAGACAGACATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1087  
1033 CGAGACCAAGCACTTCT 1089  
1088 CCTACTCAGCAACTCTCAGCTTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147  
1090 CCTACTCAGCTCTCTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACACTCTCTGAGGACACCGCTCTATGCGCTGCT 1207  
1150 TCTGGGTTCCAGGCGCTGATGCCAGGACTCCCGCAGGTTGCCCGGCTGCCCGGCTGCT 1209  
1208 ATACTGAGATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267

1210 CTACTGGCAATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269  
1268 TGTGAGCTGCTCAGGTCACATTCAGGCTTTCGAAAGCAAGCAAGCAAGCAAGCAAGCA 1327  
1270 CGGGTGTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGCT 1329  
1328 CT-----TGAAACACGAGCC 1342  
1330 CTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1389  
1343 ACCGACCTCATGATTTGCTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1402  
1390 CCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
1403 TCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462  
1450 GCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509  
1463 CCGCTTTCTTAAAGACTTAAAGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522  
1510 CCGCTTCTCAGGAGCAACCAAGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569  
1523 GCAGGAACTGATGTGAGAGATGAAAGTAGAGGATTTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
1570 GCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629  
1583 GAGGACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642  
1630 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689  
1643 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702  
1690 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749  
1703 GAGCACATTCAGAGAGAACAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1762  
1750 GACCACTTTCAAAGAGAACAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1809  
1763 GAGCATTTGAGTACAGCAACACTTGTAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1822  
1810 AAGCATTTGAAATCAGACAGCACTTTGAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1869  
1823 GGTGAGGATCAGCAGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882  
1870 GGTGAGGAGCATCGGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929  
1883 CAAGCCCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942  
1930 CAAGCTGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989  
1943 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2002  
1990 CCGCAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2049  
2003 CAACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2062  
2050 CAACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2109  
2063 CATCTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2122  
2110 TATCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2169  
2123 GATGTACTTTTGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2182  
2170 GCTGTACTTTTGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2229  
2183 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2242  
2230 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2286  
2243 TCGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2302  
2287 TGGCGGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2346





623 CCGACCCGTCGAGCAAGAAATTCACCTTAACCTTAGTTCCTTACACAGATCAAGAGCAGTAG 582  
640 GCGTCTGGGATGC-----GACCGGCCCTGGAAACCATAG 572  
683 TCGCCAGGAAGCAACGAAACCCCTGGCTTCGCCATCTCGAGGTACAAAGAGGATCTGAG 742  
673 CGTCAAGGAGCGCGGGTCCCGCTGGGCTCGCAGCCCGGGTGGAGAGGCGCGGGGG 732  
743 TCTCACAGTACAAGTGCCTTACGTTCAGTTCAGAGCCAGATGCTATCTCTCTCCGAGAGT 802  
733 CAGTGCACAGCCAAAGTCTCCGTTGCCCAAGAGGCCAGGCGTGGCGCTGCCCTGAGCC 792  
803 GGAGGAGGACCCACAGCAGGTCTACCAACCCCATCGGCAAAATCATGGGTGCCAG 862  
793 GGAGCGGACGCCGTTGGCAGGGTCTCTGGGCCACCCCGGCGAGGAGCGGTGGACCGAG 852  
863 TCTCTGCTGGTCCCGAGGTGCTTACT-----GCAGAGAAAGATTGTCTTAA 913  
853 TGACCGTGGTTCTGTGTGTGTCACTGCCAGACCCCGCAAGAAAGCCACCTCTTTGGA 912  
914 AGGAAAGTGTCTGACCTGAGTCTC-----TGGGTGGGTGTGCTGTAAACACAGCCAG 970  
913 GGTGGCTCTCTGGACCGGCCATCCACCCCATCCGTCGTCGGCGCCCGAGCACCGCGGG 972  
971 CTCACAT---CTCTGCTGTCAACACCCCGCCAAATGCCCTTTTTCAGTCTCAGGCCATTAT 1027  
973 CCCCCATCCATCGCGCCACACAGTCCCTGGGACAGCGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACGACATTTCTTTACTCAGGGGAGATGGCCAGAGGCTTAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCAGGCGA---CAAGGACAGTGTGGCGCTCCTT 1089  
1088 CTTACTCAACCACTCCAGCCTTACTGACTGGGCGCAGAGACTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCGAGCCAGCTGACTGGCGCTCGAGGCTCGTGGAGACATCTT 1149  
1148 TCTGGCTCAAGGCTTAGACATCAAGACCACTCTGCAAGCAACACCGTCTATGGGTG 1207  
1150 TCTGGGTTCCAGGCGCTGGATCCAGGAGACTCCCGCAGGTTGGCGCGCTGCCCGCAGC 1209  
1208 ATACTGGCAGATGGCGCCCTCTTCAACAGCTGCTGTGAACCATGACAGATGSCCAATA 1267  
1210 CTACTGGCAATATGGCGCCCTGTTCTGAGGCTGTTGGGAAACAGCGGCAAGTCCCTTA 1269  
1268 TGTACAGCTCTCAGTCAATGACGTTTCGACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGACACACGCC 1342  
1330 CTGTGCCGGGAGACCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAGACAGACCC 1389  
1343 ACGCACCTCATGGATTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CGTCCCTGGTGCAGTCTCCGCGACACAGCAGCCCTGGCAGGTGTACGGCTTCGT 1449  
1403 TGGGCTGTCTGCAAGGTGGTGTCTGTAGTCTCTGGGGTACCGGCAATAGAGC 1462  
1450 GCGGGCTCTGCTGCGCGGCTGGTGCCTCCCGAGGCTCTGGGGCTCCAGGCAACGAAAG 1509  
1463 CGCTTCTTTAAGAACTTAAAGATTCTATCTGTTGGGAAATACGGCAAGCTATCACT 1522  
1510 CGCTTCTCAGGACACCAAGATTCTATCTCCCTGGGAGCATGCCAAGCTCTGCT 1569  
1523 GCGGAACTGATGTGAAGATGAAGTGAAGATTGCCATCTGGCTCCGCGAGCAGCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGCACTGCGCTTGGCTGGCGAGGCCAGG 1629  
1583 GAAGGACCTGTCCCGCTGACAGCACCGTCTGAGGAGAGGATCTGGTACGTTCTCT 1642  
1630 GGTGGCTGTGTCCGGCGCAGACACCGTCTGCTGAGGAGATCTTGGCCAAAGTTCT 1689  
1643 GTTCTGCTGATGGACACATACGTGGTACAGCTGCTTAGTCACTCTTTTACATCACAGA 1702  
1690 GCACTGCTGATGAGTGTGTACGTGTCGAGGTGCTCAGGTCTTCTTTATGTACCGA 1749  
1703 GAGCACAATCCAGAAACAGGCTCTTCTTACCCGTAAGAGTGTGTGAGCAAGTGA 1762  
1750 GACCACCTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTGGAGTCHAGGCAACACTTGGAGAGTGGCTACCGGAGCTGTCAACAGAGA 1822  
1810 AAGCATTGGAATCAGACAGCACTTGAAGAGGTGCAAGTGTGCGGAGCTGTGGAAGCAG 1869  
1823 GGTCAAGCATCACAGACACCTGGGTAGCCATGCCCCATCTGCAGACTCGGCTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCAAACCGCTCGCGCCCATTTGTGAACATGATTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCCTGACGGGCTCGCGCCGATTGTGAACATGGAATAGCTCGTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAAGCAGGCCCCAGCATTTACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCCAGGCTCTCACCTCGAGGTCGAGGCACTGTTTTCAGCGTCT 2049  
2003 CAATATGAGCGGCAAAACATCTCTCACTTATGGGCTCTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGCGCGCGCCCTCTCTGGCGGCTCTGTGCTGGGCTTGACCA 2109  
2063 CATCTACAGACCTGGCGGCGCTTGTGTGCGGTGCGTCTCTGGACGACACCCAG 2122  
2110 TATCCAGAGGCGCTGGCGACCTTGTGTGCGGTGTCGGGCCCGAGACC CGCGCTGA 2169  
2123 GATCTACTTTTAAAGCAGATGTGACCGGGGCTATGATGCCATCCCCCAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGTGATGTGACGGCGGTACGACACCATCCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTTGTCCAAATATGATCAGGCACTCGGAGAGCACTGATCTGTATCCGCCAGTA 2242  
2230 CACGGAGTCTATCCCAAGCATCATCA---AACCCAGAACACGATCTGCTGCGTGGTA 2286  
2243 TGCAGTGTCTCGGAGAGATAGCAAGGCCAAAGTCCCAAGTCTCTTTAGGAGACAGGTAC 2302  
2287 TGCGTGTCTCAGAAAGCCCGCCCATGGCAGCTCCGCAAGGCCCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCAGCCATACATGGGCGAGTTCCTTTAAGCATCTCGAGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCCGTACATCGACAGTTCGTGGCTCACTGCAG-----GA 2400  
2363 TGCAGTGCATCTGAGGAACCTCCGTTGTCTATCGAGAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGATGCGGTGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCGTGTGACTTCTCTGCACTTCCTGCGTCAAGTGTGCTAAAGATTGGTGA 2482  
2461 CAGTGGGCTCTTTCAGCTCTTCTCAAGCTTCAATGCTCCACCGCCAGCGCTGCGCATCAGGGG 2520  
2483 CAGTGTCTATACGAGTCCAGGSCATCCCGCAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAGTCTCTACCTCCAGTCCAGGGGATCCCGAGGGCTCCATCTCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGGAACCAAGCTGTTTGTGTGAGGTGACCGGATGGGTT 2602  
2581 CAGCCTGTGCTACCGCCACATGGAAACCAAGCTGTTTGGGGGATTCGGCGGAGCGGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTTCTGTTGGTGGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCCTGGTTGTGTGATGATTCTTGTGTTGGTGAACCTCACTCACCCACCGCGAAAC 2700  
2663 CTTCTCAGCACCTGCTCCATGCGGTTCTCTGAGTATGGGTGATGATATAAATTTGAGAA 2722  
2701 CTTCTCAGGACCTGCTCGAGGTTCTCTGATATGCTGCTGGTGGTGAATTGGGAA 2760  
2723 GACAGTGGTGAACCTTCTCTGTGAGCGCTGGTACCCCTGGGTGGTGGAGCTCCATACAGCT 2782

2761 GACAGTGTGTAATCTCCCTGTAGAGACGAGGCGCTGGTGGCACGCTTTGTTCAGAT 2820  
2783 GCCTGCTACTGCTGTTTCCCTGTGTGGTGTGCTGTGAGACTCAGACTTTGAGGT 2842  
2821 GCGGCGCCACGCGCTATTCCCTGTGTGGGCTGTGCTGTGATACCGGACCTTGGAGT 2880  
2843 GTCTGTGCTACTACTCAGGTTATGCCACAGCTCAATTAAAGACGAGCTCACCTTCCAGAG 2902  
2881 GCAGAGCGACTACTCCAGTATGCCGAGCTCCATCAGAGCGAGTCTCACTTCAACCG 2940  
2903 TGTCTTAAGCTGGAGAGACCATCGGAAACAGCTCTGTGCTGTGGTGTGAGTG 2962  
2941 CGGCTTCAAGGCTGGAGGAAACATCGTCGCAAACTCTTTGGGCTTGGGCTGAAGTG 3000  
2963 TCACGCTCTATTCTTAGACTTGCAGTCAACAGCTCCAGACAGTCTGCATCAATATA 3022  
3001 TCACAGCTGTTCTTGGATTTCAGTGAACAGCTCCAGAGCTGTGACCAACATCTA 3060  
3023 CAAGATCTTCTGCTTCAAGGCTACAGTTCATGCAATGTGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTTCTGCTTCAAGGCTACAGTTCATGCAATGTGTGATTCAGCTTCCCTTTGA 3120  
3083 CCAGGCTGTAGAGAGACCTCACATTTCTTCTGGGATCATCTCCAGCCAGATCCTG 3142  
3121 TCAGCAAGTTTGAAGAACCACCATTTTCTGGGCTCATCTCTGACACGCGCTCCCT 3180  
3143 CTGCTATGCTATCTGATGAAGTCAAGATCCAGGAATCACATAAG-----GCCTC 3193  
3181 CTGCTACTCCATCTCTGAAGCAAGACGAGGATGTGCTGGGGCCAGGGCGCGC 3240  
3194 TGCTCTTCTCTCTGAGCGGACATGCTCTGTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTCTCGAGGCGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GCTGCTCTATCTGCTATCAAAATGCTCTGAGGCTCTGAGGACGCTCAAAACT 3313  
3301 GACTGACACCGTGTCACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GTGTGCGGAAGTCTCCAGAGGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3373  
3361 GCTGAGTGGAGTCTCCGAGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CTTAAGCAGACACTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCGCTCAGACTTCAAGACCATCTGACTGA 3454

BULT 3

-09-843-676-224

Sequence 224, Application US/09843676

Patent No. US20020164786A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: NO. US20020164786A1el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Crew LLP

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/843,676

FILING DATE: 26-Apr-2001

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US/08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US/08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US/08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTERT"

/note= "human telomerase reverse

transcriptase (hTERT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

US-09-843-676-224

Query Match 45.8%; Score 1600.8; DB 10; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGCGCTTGAGCACATGACCGCGCTCTCTGCTGCGCGGCTGCTCTCTGCTGCG 82  
DB 40 CCGCGCACCGCGCGATGCGCGCTCCCGCTCCGAGCGCTGCGCTCTCTGCTGCG 99  
QY 83 CAGCCGATACCGGAGGTGTGGCGCTTGTGCGCGCGCTGCGCGCGCGCGCGG 142  
DB 100 CAGCCACTACCGGAGGTGTGGCGCTTGTGCGCGCGCTGCGCGCGCGCGCGG 159  
QY 143 CAGCGCGCTTGTGCAACCGCGGACCGGAGATCTACCGCATTTGGTTGCCAATGCT 202  
DB 160 CTGCGCGCTGCTGCGCGCGCGGACCGCGCGCTTTCGCGCGCTGCTGCGCGCT 219  
QY 203 AGTGTGATGCTGCGGCTTCACAGCTCCAGCTCCAGCTGCGCGCTTCTCTCCACAGGTGTC 262  
DB 220 GGTGTGCTGCTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTC 279  
QY 263 ATCCCTGAAGAGCTGTGGCGAGGTGTGAGAGACTCTGCGAGCGCGCAACGAGAGAA 322  
DB 280 CTGCTGAAGAGCTGTGGCGCGGAGTGTGCTGAGAGGCTGTGCGAGCGCGCGCGG 339  
QY 323 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382  
DB 340 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
QY 383 CACTAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442  
DB 400 CACCCAGCGTGTGCGAGCTTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGG 459  
QY 443 TGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502  
DB 460 GCGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519

503 CTGTGCTTTATCTTCTGGTCCGCCAGCTGTGCTTACAGAGTGTGTGCTTCCCT 562  
520 CTGCGGCTCTTTGTGCTGGTGGCTCCAGCTGCGCTTACCAGGTGTGCGGCGCGCT 579  
563 GTACCAAAATTTGTGCCACCAAGATATCTGGCCCTCTGTGCTGCTAGTTACAGGCCAC 622  
580 GTACAGCTCGGCTGCCACTCAGGCCGCGCCCGGCCACACGCTAGTGACCCCAAG 639  
623 CGACCCGCTGGGAGGAATTCATAACCTTAGTTCCTTAACAAGATCAAGACAGTAG 682  
640 GCCTCTGGGATGC-----GAACGGGCTTGGAAACCATAG 672  
683 TGCCAGGAGACACCGAACCCCTGGCTTGCATCTCGAGGTACAAGAGGCATCTGAG 742  
673 CGTACGAGAGCGCGGCTCCCTCGGCTCGACCGCGGCTGGAGAGCGCGGGG 732  
743 TCTCACAGTACAAGTGTGCTTCAAGTCAAGAGCCAGATGCTATCTGTCCGAGAGT 802  
733 CAGTGCACCGGAAGTCTGCGTTCGCCAAGAGGCCAGGCGTGGCGCTGCCCTGAGCC 792  
803 GAGGAGGAGACCCACAGCAGGTGCTACCAACCCATCAGGCAATCATGGTGCACAG 862  
793 GGAGCGGACGCGCTTGGGACAGGGTCTTGGGCCACCCCGGCGAGGACGCTGGACCGAG 852  
863 TCCTGCTCGGTCGCCCGAGGTGCTACT-----GCAGAGAAAGATTTGTCTTCTAA 913  
853 TGACCGTGGTTCTGTGTGTGTCACTGCCAGACCCGCGCAAGAGCCACTCTTTTGA 912  
914 AGAAAGGTGTCTGACCTGAGTCTCTC---TGGTGGTGTGTCTGTAAACAGACCCAG 970  
913 GGTGGGCTCTTGGCACCGGCACTCCACCCATCCGTTGGGCGCCAGCACACGCGGG 972  
971 CTCACAT---CTCTGCTCTACCAACCCCGCCAAATGCTTTTCTGCTCAGGCTCAGGCCATTTAT 1027  
973 CCCCCATCCACATCGGCGCAACAGTCCCTGSGACAGGCTTGTCCCGGCTGACGC 1032  
1028 TGAGACAGACATTTCTTACTCCAGGGAGATGGCCAGAGGCTTAAACCCCTCATTT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAAGGCA---CAAGAGCAGCTGCGGCGCTCTT 1089  
1088 CTTACTCAGCAACCTCCAGCTAACTTGACTGGGCGCAGGAGACTTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCCACGCTGACTGCGCTCGAGGCTCTGTGAGACCATCTT 1149  
1148 TCTGGGCTCAGGCTTAGACATCAGGACCACTCTGAGGACACACCGTCTATCGGTG 1207  
1150 TCTGGGTTCCAGSCCTGGATGCCAGGACTCCCGCGAGTTGCGCGCTGCGCCAGCG 1209  
1208 ATACTGGCAGATCGGCGCTGTTCCACAGCTGCTGTGAACCATGACAGTGCRAATA 1367  
1210 CTACTGGCAATCGGCGCTGTTTCTGAGCTGTTGGAAACACGCGAGTGCCTTA 1269  
1268 TGTGAGATCTCTCAGTCTCATTTGCAAGTTTCGAACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACCACTGCGCGCTGCGAGCTGCGGTACCCAGCAGCGGTGT 1329  
1328 CT-----TGACACAGCGCC 1342  
1330 CTGTGCCGGAGAACCCCGAGGGTCTGTGGCGCGCCCGGAGGAGGAGACACAGACCC 1389  
1343 ACCGCACTCATGGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402  
1390 CGGTGCGCTGTCAGCTGCTCGCGCAGACACAGACGCCCTGGCAGGTGTACGGTTCTGT 1449  
1403 TCGGCGCTGCTCTGAAAGTGTGTCTGATCTCTGGGGTACACAGGACATGAGCG 1462  
1450 GCGGCGCTGCTCGCGCGGTGGTGGCGCCAGGCTCTGGGGCTCCAGGCAACGAACG 1509  
1463 CCGTCTCTTAAAGATTTATCTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGAAACACCAAGAGTTTATCTCCCTGGGAGACATGCAAGCTCTCGCT 1569

1523 GCAGAACTGATCTGGAGATGAAGTAGAGATTGCCACTGGCTCGCAGAGCCCGG 1582  
1570 GCAGAGCTGACGTGGAGATGAGCTGCGGAGCTGCGTTGGCTGCGCAGGAGCCAGG 1629  
1583 GAAAGACCGTGTCCCGCTGCAGAGCACCGCTCTGAGGAGAGATCTGCTACGTTCTCT 1642  
1630 GTTGGCTGTGTTCGCGCGCAGAGCACCGCTCTGCTGAGAGATCTTGGCCCAAGTTCT 1689  
1643 GTTCTGGCTGATGACACATACGTGCTGACAGTCTTAGGTCTATTTTATCATCAGA 1702  
1690 GCATCTGGCTGATGATGTGTAAGTCTGCTGAGCTGCTCAGGTCTTTCTTTATGTACGA 1749  
1703 GAGCACAATCCAGAGAACAGGCTCTTTCTACCGTAAGAGTGTGAGCAAGTGTGA 1762  
1750 GACCACTTTTCAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACCTTGAAGAGTGTGGCTACGGGAGCTGTCAACAGGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGCACTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTACGGCATCACAGAGCACCTGGCTAGCCATCTGCAGACTGCGCTTCAATCCC 1882  
1870 GGTACGGCAGCATCGGAGAGCCAGCCGCTGCTGACGTCCAGACTCGCTTCAATCCC 1929  
1883 CAGCCCAACCGCTGCGGCCATTTGTAAATGATGATGATGATGATGATGATGATGAT 1942  
1930 CAGGCTGACGGCTGCGGCCATTTGTAAATGATGATGATGATGATGATGATGATGAT 1989  
1943 GGGCAGAGAGAGCAGCCAGCATTTTACCCAGCTCTCAAGCTCTCTTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGCGCTCTACCTCAGGTTGAGGCACTGTTTTCAGCGTCT 2049  
2003 CAATATGAGCGGCAAAACATCTCTTATGGGTCTTCTGATGATGATGATGATGATGAT 2062  
2050 CAATACAGCGGCGCGGCCCTCTGGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2109  
2063 CATCTACAGGACCTGCGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122  
2110 TATCCAGAGGCTGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169  
2123 GATGTACTTTTGAAGCAGATGTGACCGGGCTATGATGCCATCCCGAGGCTAAGCT 2182  
2170 GCTGTACTTTTGAAGTGGATGTGACGGCGGTACGACACATCCCGAGGAGCT 2229  
2183 GGTGAGTGTGCTGCAATATGATCAGGCTCTGAGAGACGCTGCTGCTGCTGCTGCTGCTG 2242  
2230 CAGGAGGTATCGCCAGCATCATCA---AACCCCAACACACGCTGCTGCTGCTGCTGCTG 2286  
2243 TGCACTGTCTCGGAGAGATAGCCAAAGCCAAAGTCCCAAGTCTTTTAGGAGACAGCTCAC 2302  
2287 TGCGGTGTCAGAGCGCGCCATGGGCGCTCGGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CAGCTCTCTGACCTCCAGGCATACATGCGGCGAGTCTTCTTAAAGCTCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGCTACATCGACACATCTGCTGCTCACTGCTGCTGCTGCTG 2400  
2363 TGCCAGTCTCATGAGGACTCCGTTGCTATCGAGCAGCATCTCTATGATGAGAGCAG 2422  
2401 GACCAAGCGCTGAGGAGATGCGGCTGCTATCGAGAGAGCTCTCTCTGATGAGGCGCAG 2460  
2423 CAGCAGCTGTTTGAATCTTCTTCTGCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTCGAGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGTGTCTATGAGGCTGCGAGGCTATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2542  
2521 CAAGTCTTCTGCTGAGTGTGAGGAGACAGCTGTTTTCGCGGAGTTCGGCGGAGCGGCT 2580  
2543 CAGTCTGTGTTTTCGAGACATGAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2602  
2581 CAGCTGTGCTTACGCGACATGAGGAGACAGCTGTTTTCGCGGAGTTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTTTGTGATGACTTTTCTGTTGGTGACGCTCATCTTGGACCAACCAAAAC 2662

2641 GCTCTGCGTGTGGTATGTTCTTGTGTGACACCTCACCTCACCGGAAC 2700  
2663 CTTCTCAGACCCCTGGTCCATGGGTTCTGATGATGGGTGCAATGAATGAGAA 2722  
2701 CTTCTCAGACCCCTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGAATGCGGAA 2760  
2723 GACAGTGTGAACCTCCCTGTGGAGCCTGTACCTGGGTGTGAGCTCCATACAGCT 2782  
2761 GACAGTGTGAACCTCCCTGTAGAGACGAGCCCTGGGTGACGCGCTTTGTCAGAT 2820  
2793 GCCTGCTCAGTCCCTGTTTCCCTGGTGTGCTGCTGAGCACTCAGACTTTGAGGT 2842  
2821 GCGGCCCCACGCGCTATTCCCTGGTGTGCGGCTGCTGCTGATACCGGACCCCTGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTATGATCCGAGACCTCAATTAAGCAGGCTCACCTCCAGAG 2902  
2881 GCAGAGCAGCTACTCCAGCTATGCCGAGCTCCATCAGAGCAGCTCACCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAAGACCATCGGAACAGCTCCTGTGCGTCTTGGGTTGAAGTG 2962  
2941 CGGCTTCAAGGCTGGGAGGAACATGCGTGGCAACTCTTTGGGCTCTTGGGCTGAAGTG 3000  
2963 TCACGCTCTATTCTAGACTTGCAGGTGAACAGCCTCCAGACAGCTGCTGATATATA 3022  
3001 TCACAGCTGTTTCTGGATTTGCAAGTTGAACAGCCTCCAGACGCTGTGCAACCACTA 3060  
3023 CAAGATCTTCTGCTTCAAGGCTACAGGTTCCATGCTATGTGATTCAGCTTCCCTTGA 3082  
3061 CAAGATCTTCTGCTGAGGCTGACAGGTTTCAAGCTATGTGCTGAGCTCCATTTCA 3120  
3083 CGAGGCTGTTAGGAAGAACCTCACTTCTTGTGGGATCATCTCCAGCAAGCAATCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGCGCTCCCT 3180  
3143 CTGCTATCTATCTGAGGTCAGGTCAGGATCCAGGAATGACATAAG-----GCCTC 3193  
3181 CTGCTACTTCTTCTGAGGTCAGGTCAGGATGTCGCTGGGCGCCAGGGCGCCGC 3240  
3194 TGGCTCTTCTCTGAGGCGGACATTTGGCTCTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTGCGCTCGAGGCGGTGAGTGGTGTGCGCACCAAGCATCTCTGCTCAAGCT 3300  
3254 GCTGCTCACTTCTGCTATCAATAATGTCTTGGGACCTCTGAGAGAGCCCAAAAT 3313  
3301 GACTGACACCGTGTCACTACCTACCTGCGGCTCCTGGGTCACTCAGAGACGCCAGCA 3360  
3314 GCTGTGCGGAGCTCCAGAGGCGCAATGACCATCTTAAAGCTGACGCTGACCCAGC 3373  
3361 GCTGAGTCGAGGCTCCCGGAGAGAGCTGACTGCTGAGGCGCGAGCCMACCCGCG 3420  
3374 CTAAGCAGAGCTTTCAGACCTTTTGGACTAA 3407  
3421 ACTGCGCTCAGACTTCAAGACCATCTCTGGACTGA 3454

ULT 4

09-953-052-1

Sequence 1, Application US/09953052

Patent No. US20020173476A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin B.

Andrews, William H.

TITLE OF INVENTION: Antisense Compositions for Detecting and

Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/953,052  
FILING DATE: 14-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/052,919  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "human telomerase reverse  
transcriptase (hTRT)"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-953-052-1

Query Match

Best Local Similarity

Matches 23/9; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

Qy	23	CCCGGCGCTTTGAGCAATGACCCGCGCTCCTCGTTGCCCCCGGTCGCTCTCTGCTCGG	82
Db	40	CCGGCCACCCCGCGATGCGCGCTCCCGCTGCCGAGCGCTCCTCTGCTCGG	99
Qy	83	CAGCCGATACCGGAGGTGTGCGCGCTGCGCACTTTGTGCGGCGCTGGGGCCCGAGGG	142
Db	100	CAGCCACTACCGGAGGTGCTCGCGCTGCGCACTTTGTGCGGCGCTGGGGCCCGAGGG	159



2287 TCCGCTGCTCCAGAGGCGCGCCATCGGCAAGGCGCTTCAAGAGCCACAGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGGACATACATGCGCAGTCTCTTAAGCATCTCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATCGAAGTTCTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGTGCTAGAGAACTCCGTTGTTCATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATCGCGTCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCCTGTTTGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGCCTCTCTCGACGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGGTGCTATAGCAGTCCAGGCGCATCCCGAGGCTCCAGGCTTCCAGCTTCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCCAGTCCAGGCGATCCCGAGGCTCCAGGCTTCCAGCTTCCAGCTGCTCTG 2580  
2543 CAGTCTGTTTCCGAGACATGAGAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2602  
2581 CAGCCTGTGCTACGCGGACATGAGAACTGTTTGGGGGATTCGGCGGACGCGCT 2640  
2603 GCTTTTACGCTTGTGTGATGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2662  
2641 GCTCTGCTTGTGTGATGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2663 CTCTCTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
2701 CTCTCTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2723 GACGTGTGAACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782  
2761 GACGTGTGAACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCCCCAGGCGCTTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2902  
2881 GCAGAGCAGCTACTCCAGTATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
2903 TGTCTTAAAGCTGGGAGACCATGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGGCTGGGAGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
2963 TCACGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3022  
3001 TCACAGCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060  
3023 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082  
3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3083 CAGCGTGTGAGGAGAACCTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
3121 TCAGCAAGTGTGAGGAGAACCTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
3143 CTGCTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3193  
3181 CTGCTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240  
3194 TGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
3241 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3313  
3301 GACTCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCGGAGCTCCAGAGGCGGCAATGACCATCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 3373  
3361 GCTGAGTCCGAGCTCCGCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420

QY 3374 CCTAAGCAGCAGCTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

## RESULT 5

US-10-053-758-224  
; Sequence 224, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; ; Lingner, Joachim  
; ; Nakamura, Toru  
; ; Chapman, Karen B.  
; ; Morin, Gregg B.  
; ; Harley, Calvin  
; ; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1e1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-0029300S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 56..3454  
; OTHER INFORMATION: /product= "hTRT"  
; /note= "human telomerase reverse  
; transcriptase (hTRT) catalytic protein  
; component"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-10-053-758-224

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;



Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGGGCTTGGACACAAATGACCCGGCTCTCTGTTGCCCGCGGTGGCTCTCTGCTGGC 82  
|||||  
40 CCGGGCCACCCCGGATGCCCGCGCTCCCGCTGCGGAGCCGCTGCGCTCCCTGCTGGC 99  
|||||  
83 CAGCCGATACCCGGAGGTGTGCGCGCTGCGAACCTTTGTGGCGCGCTGGGGCCCGAGGG 142  
|||||  
100 CAGCCACTACCCGAGGTGTGCGCGCTGCGCACTTTGTGGCGCGCTGGGGCCCGAGGG 159  
|||||  
143 CAGCGCGCTTGTGCAACCCGGGACCCGAGATCTACCGCACTTTGTGGTGGCCAAATGCTT 202  
|||||  
160 CTGGCGCTGTGTGACGCGCGGGACCCGCGGCTTTCCGCGCGCTGTGTGGCCAGTGCCT 219  
|||||  
203 AGTGTGATGCACTGGGGCTCACAGCTCTCACTGCGGACCTTTCTCTTCCACAGGTGTC 262  
|||||  
220 GGTGTGCTGCTTGGGACGACAGCGCGCGCCCGCCCGCTCTCTTCCGCGAGGTGTC 279  
|||||  
263 ATCCCTGAAGAGCTGTGGCCAGGGTTGTGCAGAGACTCTGCAGGCGCAACGAGAGAAA 322  
|||||  
280 CTGCTTGAAGGAGCTGTGTCGCCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAAGAA 339  
|||||  
323 CGTGTGCTTTTGGCTTTGAGCTGCTTAAAGAGCCAGAGCGGGCTCCCATGGCCTT 382  
|||||  
340 CGTGTGGCTTTGCGCTTCCGCTGCTGACAGGGCCCGCGGGCGCCCGGAGGCTT 399  
|||||  
383 CACTAGTAGCTGCGTAGCTACTTTCGCCAACACTGTATTAGAGACCTGTGCTGTCAGTGG 442  
|||||  
400 CACCACCGCTGGCAGCTACTCTGCCCAACAGGGTGACCGACACTGTGGGGGAGCGG 459  
|||||  
443 TGCATGGATGCTACTGTGTGAGCCGAGTGGGCGACGACCTGTGTGTCTACTGTGGCACA 502  
|||||  
460 GCGTGGGGCTGCTGCTGCGCGCGCTGGGGGACACGCTGTGTGTCTACCTGTGGCAGC 519  
|||||  
503 CTGTGCTCTTTATCTTCTGGTCCCCCAGCTGTGCTTACCAGGTGTGGGTCTCCCTT 562  
|||||  
520 CTGCGGCTCTTTGCTGTGTGCTTCCAGCTGTGCGCTTACAGGTGTGGGGCGCGCT 579  
|||||  
563 GTACCAATTTGTGCCACACAGGATATGCGCCCTCTGTGTGCGCTAGTTACAGGCCAC 632  
|||||  
580 GTACCAAGCTCGGCGTGGCACTCAGCGCGCGCGCCCGCCACACGCTAGTGGACCCGAG 639  
|||||  
623 CCGACCCGTGGCAGGAATTTCACTAACTTAGTGTCTTACACAGATCAAGAGCAGTAG 682  
|||||  
640 GGTCTTGGGATGC-----GAAAGGGCTTGGAAACCATAG 672  
|||||  
683 TGCCAGGAAGCACCGAAACCCCTGCGCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 742  
|||||  
673 CGTCAAGGAGCGGGGTCCCTGCGCTGCGACCCCGGCTGCGAGGAGCGCGGGG 732  
|||||  
743 TCTCACAGTACAGTGTGCTTCACTAAGAGGCGAGATGCTATCTGTCCGAGAGT 802  
|||||  
733 CAGTGCACCGAAGTCTGCGCTTGCCTAAGAGGCGCCAGGCGTGGCGTGCCTCTGAGCC 792  
|||||  
803 GAGGAGGACCCCAAGGAGGTGTACTCAACACCCCATCAGGCAAAATCATGGTGCCAAAG 862  
|||||  
793 GAGCGGACGCGCGTGTGGCAGGGTCTGCGCCACCCCGGCGAGGACGCTGGAACCGAG 852  
|||||  
863 TCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
|||||  
853 TGACCGTGTGTTCTGGA 912  
|||||  
914 AGGAAAGGTGTGTGACCTGAGTCTCTC-----TGCGTGGGTGTGTGTGTGTGTGTGTGTGT 970  
|||||  
913 GGTGTGCTCTCTGTGGACGCGCACTCCACCCCATCTCGTGGCGCGCAGCACACGCGGG 972  
|||||  
971 CTCACAT---CTCTGCTGTACAGCCCGCGCAAAATGCTTTTCTGCTGAGGCGCATTTAT 1027  
|||||  
973 CCCCCATCCATCCGCGCCACCAAGTCTCTTGGGACACGCTTGTCCCGCGGTGTACGC 1032  
|||||  
1028 TGAGACACAGATTTCTTTTCTTCCAGGGGAGATGGCCAGAGGCTCTTAAACCCCTCAT 1087  
|||||  
1033 CGAGACCAAGCACTTCTCTTCTCTCAGGCGA---CAAGGAGAGCTGGCGGCTCTT 1089  
|||||

QY 1088 CTTACTCAGCAACCTCCAGCCTTAACCTTGAATGCTGGGCGCAGGAGCTGGTGAGATCATCTT 1147  
|||||  
Db 1090 CTTACTCAGCTCTCTGAGGCGCCAGCCTGACTGGCGCTCGGAGGCTCGTGAGACCACTTT 1149  
|||||  
QY 1148 TCTGGGTCAAGGCTTAGGACATCAGGACCACTCTGCGAGACACACCGCTTATCGCGTGG 1207  
|||||  
Db 1150 TCTGGTTCCAGGCGCTTGGATGCCAGGACTTCCCGCAGGTTGCCCGCTGCCCGCAGCG 1209  
|||||  
QY 1208 ATACTGGCAGATGGCGGCTTGTTCGAAAGCTGTGGTGAACCATGCGAGAGTGGCAATA 1267  
|||||  
Db 1210 CTACTGGCAATGGCGGCTTGTTCGAGCTGCTTGGGAACCGCGCAGTGGCGCCCTA 1269  
|||||  
QY 1268 TGTGAGACTCTCTCAGGTCAATGTCAGGTTTGAAGAGCAACCAACAGGTGACAGATGC 1327  
|||||  
Db 1270 CGGGTGTCTCTCAGAGCGCACTGCCCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
|||||  
QY 1328 CT-----TGAAACCAAGGCC 1342  
|||||  
Db 1330 CTGTGCCCGGAGAACCCAGGCTCTGTGGCGGCGCCCGAGGAGGAGGACACAGACCC 1389  
|||||  
QY 1343 ACCGACCTCATGATTTGCTCCGCTGACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
|||||  
Db 1390 CGGTGCGCTGTGAGCTGTCTCGCCAGCAGCAGCGCTTGCAGGTGTACGGCTTCT 1449  
|||||  
QY 1403 TCGGGCTGTCTCTGCAAGGTGTGTCTGTCTGTCTGTGGGTACCGAGCACAATGAGCG 1462  
|||||  
Db 1450 CCGGCGCTGCTGCGCGGCTGTGTGCGCCAGGCTCTGGGGCTCCAGGCAACCAAGCG 1509  
|||||  
QY 1463 CGGTCTCTTAAGAACTTAAGAGTTTCACTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
|||||  
Db 1510 CGGTCTCTCAGGAACACCAAGAGTTTCACTCTCCCTGGGGAACATGCAAGCTCTCGT 1569  
|||||  
QY 1523 GCAGAACTGATGTGGAGATGAAAGTAGAGATTGCCACTGGCTCCGAGCAGCGCGGG 1582  
|||||  
Db 1570 GCAGAGCTGACGTGGAAGATGAGCTGCGGAGTGGCTTGGCTGCGCAGGAGCGCAGG 1629  
|||||  
QY 1583 GAAGGACCGTGTCCCGCTGCGAGAGCAACCGTCTGAGGAGAGGATCTTGGCTACGTTCCT 1642  
|||||  
Db 1630 GGTGTGGTGTCTTCCGCGCAGAGCAACCGTCTGCGTGAAGGATCTTGGCAGGTTCCT 1689  
|||||  
QY 1643 GTTCTGGCTGATGACACATACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1702  
|||||  
Db 1690 GCATGT 1749  
|||||  
QY 1703 GAGCACTTCAGGAAGAACAGGCTCTTCTTCTACCTGTAAGAGGTGTGTGTGTGTGTGTGT 1762  
|||||  
Db 1750 GACCACTTTCAAAAGAACAGGCTCTTCTTCTACCGGAAGAGTGTCTGAGCAAGTTGCA 1809  
|||||  
QY 1763 GAGCACTTGGAGTCAAGCAACACCTTGAGAGAGTGGCTACGGAGCTGTCAAGAGGA 1822  
|||||  
Db 1810 AAGCACTTGGAAATCAGACAGCACTTGAAGAGGTGTGAGCTGCGGAGCTGTTCGAGAGCAG 1869  
|||||  
QY 1823 GGTCAAGCATCAACAGGACACCTGGGTAGCCATGCCCATCTGCGAGCTGCGCTTCAATCC 1882  
|||||  
Db 1870 GGTCAAGCAGCATCGGAAGCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929  
|||||  
QY 1883 CAGGCCAAAGCGCTGCGGCCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
|||||  
Db 1930 CAAGCTTACCGGGCTGCGGCGGATTTGGAACATGGAACCTGCTGCGGAGCCAGAACGTT 1989  
|||||  
QY 1943 GGGCAAGAGGAGGAGGCGCCAGCAATTTCAACCGAGCTCTCAAGACTCTCTTTCAGCATGT 2002  
|||||  
Db 1990 CCGCAGAGAAAGAGGCGCGAGCGCTCTCACTCGAGGTGAAGGCACTGTTCAGCGTGT 2049  
|||||  
QY 2003 CAATATGAGGAGCAAAACATCTCCTACCTTATGGGTCTTCTGTACTGGGTATGATGA 2062  
|||||  
Db 2050 CAATCTACAGCGGCGCGGCGCCCGCGCTCTTGGGCGCTCTGTGTGTGTGTGTGTGTGTGT 2109  
|||||  
QY 2063 CATCTACAGGACCTTGGGGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
|||||  
Db 2110 TATCCACAGGCGCTTGGCGCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
|||||









2401 GACAGCCCGCTGAGGATGCGTGTCTGACGACAGAGCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCGCTGTTGACCTTCTCGCTGCACTTCTCGCTGCAAGTGTGTAAGATTGGTGA 2482  
2461 CAGTGGCGCTCTTCGACGCTCTCTAGCTTCTATGTGCCACCGCGTCCGATCAGGGG 2520  
2483 CAGTGTCTATACGACGATGCGAGGATCCCGAGGCTCCAGCTATCCACCTCTCTG 2542  
2521 CAAGTCTCTGTCAGTGCAGGGATCCCGAGGCTCCATCTCTCCACGCTCTCTG 2580  
2543 CAGTCTGTTTCCGAGACATGAGAACTGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCCTGTGTACGGCAGATGAGAACTGCTGTTGCGGATTTCCGCGGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGACATTTCTGTGTGAGCTCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCCTCGGTTGTGATGATTTCTGTTGGTACCTTCACTTCACTTCACTTCACTT 2700  
2663 CTTCCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
2701 CTTCCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2723 GACAGTGTGACTTCCCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782  
2761 GACAGTGTGACTTCCCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCCCCAGGCTATTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGTGACTCTCAGGTTATGCCAGACCTCAATTAAGACAGCTTCACTTCCAGAG 2902  
2881 GCAGAGGACTACTCCAGCTATGCCGAGCTTCCATCAGAGCAGCTTCACTTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAGACCATGCGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGGGAGAACATGCTGCAACTCTTTGGGGCTTGGGCTGAAAGT 3000  
2963 TCAGGCTTATTTCTAGCTTCAAGTGAACAGCTTCCAGACAGCTGCTGCTCAATATATA 3022  
3001 TCAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060  
3023 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082  
3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3083 CCAGCGTGTAGGAAGAACTCAGATTTCTTCTGGGCATCATCTCCAGCAAGCATCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCTCTGACAGCGCTTCCCT 3180  
3143 CTGCTATGCTATCTGAGGTCAAGATCCAGGATGACACTAAAG-----GCCTC 3193  
3181 CTGCTACTCTCCTGAAGCCCAAGAACCGAGGATGCTGCTGGGGGCCAAGGGCGCGC 3240  
3194 TGGCTCTCTTCTCTCTGAGCCGCCACATTTGCTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGCTGCTCTATCTGCTATCAAAATGCTTCTGGAACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTCGACACCGTGTCACTTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGTGTGAGTGTGAGTGTGAG 3373  
3361 GCTGAGTGTGAAGCTCCGAGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CCTAAGCAGACTTTTCAAGACATTTTGAAGTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTGAGACTGA 3454

US-10-054-611-224  
; Sequence 224, Application US/10054611  
; Publication No. US20030059787A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. US20030059787A1Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/054,611  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/854,050  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 56..3454  
; OTHER INFORMATION: /product= "hTERT"  
; /note= "human telomerase reverse  
; transcriptase (hTERT) catalytic protein  
; component"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-10-054-611-224

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 23/9; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCCGCCCTTGAGCAATGACCGCGCTCCCTGCTGCGCGGTGGCGCTGCTGCTGCG 82  
DB 40 CCCGCCACCCCGCGATGCCGCGCTCCCGCTGCGAGCCGCTGCTGCTGCG 99  
QY 83 CAGCGATACCGGAGGTGTGCGCTGGCAACTTTGTGCGGCGCTGGGCGCCGAGGG 142  
DB 100 CAGCCACTACCGGAGGTGTGCGCTGGCACCTTGTGCGGCGCTGGGCGCCCGAGGG 159



143 CAGGGCTTGTGCAACCCGGGACCCGAGATCTACCGACATTTGGTTGCCAATGCC 202  
160 CTGGCGCTGTGTGAGCGCGGGACCCGGGGCTTTCCGGCGCTGGTGGCCAGTGCT 219  
203 AGTGTGATGACACTGGGGCTCAGAGCTCCACTGCCGACCTTTCTTCCACAGGTGTC 262  
220 GGTGTGCTGCTTGGGACGACCGCGCCGCCCGCCCGCCCTCTCTTCCGCGAGGTGTC 279  
263 ATCCCTGAAGAGCTGTGGGTCAGGGTGTGCGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCTGAAGAGCTGTGGGCCCGAGTGTGCGAGGCTGTGCGAGCGCGGGCGAAGAA 339  
323 CGTCTGGCTTTTGGCTTTGAGCTGCTTAAGAGGCCAGAGCGGGCCCTCCATGGCCTT 382  
340 GTGTGCGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 399  
383 CACTAGTAGCGTGTAGTACTTTGCCCAACACTGTTATTGAGACCCCTCGGTGTAGTGG 442  
400 CACCACAGCGTGCAGTACTCTGCCCAACACAGGTGACCGACGCACTGCGGGGAGCG 459  
443 TGCATGATGCTACTGTGAGCGAGTGGGCGAGGAGCTGCTGCTACTGCTGTCACCA 502  
460 GCGTGGGGGTGTGTGCGCGCGCTGGGCGAGCGAGTGTGCTGCTGCTGCTGTCACG 519  
503 CTGTGCTCTTTATCTTGTGTGCCCCCAGCTGTGCTACAGGTGTGTGGGTCTCCCT 562  
520 CTGCGGCTCTTTGTGTGTGTGCTCCAGCTGCGCTTACAGGTGTGCGGGCGCGCT 579  
563 GTACCAATTTGTGCGACAGGATATCTGCGCTCTGTGTGCTGCTGCTGCTGCTGCTG 622  
580 GTACAGCTGCGCGTGTGCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG 639  
623 CGAGCCCTGCGGCGAGGATTTCACTAATTTAGTCTTCTTACACAGATCAAGAGCAGTAG 682  
640 GGTCTGGGATGC-----GAGCGGCTGGAACTAG 672  
683 TGGCAGGAAGCAACCGAAACCCCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 742  
673 CGTACGGAGGCGGGGTCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 732  
743 TCTACAGTACAGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 802  
733 CAGTCCAGCGGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 792  
803 GGAGGAGGAGCCCAAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 862  
793 GGAGGAGGAGCCGCTTGGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 852  
863 TCCTGCTCGTTCCTGAGT 913  
853 TGACCGT 912  
914 AGGAAAGGT 970  
913 GGT 972  
971 CTCACAT---CTGT 1027  
973 CCCCCATCCCATCGCGGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032  
1028 TGAGACAGACATTTCTTGT 1087  
1033 CGAGACAGACATTTCTTGT 1089  
1088 CTTACTAGCAACCTTCCAGCTTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147  
1090 CTTACTAGCTCTGT 1149  
1148 TGTGGCTTCAAGGCTTGT 1207  
1150 TGTGGTTCAGGCGCTTGT 1209

QY 1208 ATACTGGAGATGGGCGCCCTGTTCACACAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1267  
Db 1210 CTACTGGCAATCGGCGCCCTGTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269  
QY 1268 TGTGAGACTCTCAGGTACATTTGAGGTTCGAAACAGCAACAAACAGGTGTGACAGTGC 1327  
Db 1270 CGGGTGTCTCTCAAGACGCACTGCCGCTGTGAGTGGGTCACTCCAGCAGCGGTGT 1329  
QY 1328 CT-----TGAAACACAGGCC 1342  
Db 1330 CTGTGCGCGGAGAACCCAGGGCTCTGTGCGGCGCCCGAGGAGGAGGACACAGACCC 1389  
QY 1343 ACCGCACTCATGATTTGCTCGCTCGCTGACAGAGTCCCTGCGAGGTATATGTGTTCCT 1402  
Db 1390 CGTGTGCTGTGTGAGTGTCTGCGCAGCAGAGCCCTGCGAGGTGTACGGTTTGT 1449  
QY 1403 TCGGGCTGTCTCTGCAAGGT 1462  
Db 1450 GCGGGCTGTCTGCGCGGCTGTGCGCCAGGCTCTGTGGGCTCCAGGCAACAGGACG 1509  
QY 1463 CGCTTCTTTAAGAACTTAAGAACTTCACTCTGTTGGGAAATACGCAAGCTATCACT 1522  
Db 1510 CGCTTCTCTCAGGAAACCAAGAACTTCACTCTGTTGGGAAATACGCAAGCTATCACT 1569  
QY 1523 GCAGGAATCTGATGTGAAGATGAAGATGAGGATTCCTGCTCCGAGAGCCCGGG 1582  
Db 1570 GCAGGAGCTGACGTGTGAAGATGAGGATGCGGCTGCGCTGCGCAGGAGCCAGG 1629  
QY 1583 GAAGGACGCTGCTCCGCTGCGAGAGCCGCTGTGAGGAGAGGATTCCTGCTACGTTCT 1642  
Db 1630 GGTGTGTGTGTTCGCGCGCAGAGCCGCTGTGCTGAGGAGATCCTGCGCAAGTTCT 1689  
QY 1643 GTTCTGCTGATGACACATAGT 1702  
Db 1690 GCAGT 1749  
QY 1703 GAGCAATTCAGAGAAACAGGCTCTTCTTACGTAAGAGTGTGTGTGTGTGTGTGTGTGT 1762  
Db 1750 GACCACTTTCAAAAGAACAGGCTCTTTTCTACGGAAGAGTGTGTGTGTGTGTGTGTGT 1809  
QY 1763 GAGCATTTGAGTCAGCAACACCTTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1822  
Db 1810 AAGCATTTGAAATCAGACAGCATTTGAGAGGTTGAGCTGTGCGGAGCTGTGTGTGTGTGT 1869  
QY 1823 GGTGAGCATTCACAGGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1882  
Db 1870 GGTGAGGACATCGGGAAGCAGGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1929  
QY 1883 CAAGCCCAAGCGCTCGGCGCCATTTGTGAACATGATGATGATGATGATGATGATGATGATGAT 1942  
Db 1930 CAAGCTGTGAGGCTCGGCGCGGATTTGTGAACATGATGATGATGATGATGATGATGATGAT 1989  
QY 1943 GGGCAGAGGAAGCAGGCGCCAGCATTTACCCAGGCTCTCAAGACTCTCTTTCAGCATGCT 2002  
Db 1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACTCGAGGGTGAAGGCACTGTTTCAGCGTGTCT 2049  
QY 2003 CAATATGAGCGCAAAACATCTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2062  
Db 2050 CAATACAGCGCGCGCGCGCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2109  
QY 2063 CATCTACAGACCTGCGCGGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
Db 2110 TATCCACAGGCTGTGCGGCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
QY 2123 GATGTATTTGTTAAGGCGAGATGTGACCGGGGCTTATGATGCTTCCCTCCAGGTTAAGCT 2182  
Db 2170 GCTGTATTTTGTCAAGTGTGATGTGACGGGCGCGTACGACACCTTCCCTCCAGGAGCT 2229  
QY 2183 GGTGTGAGT 2242  
Db 2230 CAGGAGGTCTATCGCAGCATCTCA---AACCCAGAACAGTACTGCTGTGTGTGTGTGTGT 2286  
QY 2243 TGCAGT 2302

2287 TGCCGTGTCAGAGGCGCCGATGCGCAGCTCCGAGAGCCCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTTCAGCCATACATGAGGCGAGTCTCTTAAAGCATCTGAGAGATTGAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGTGGCTCACCTGCAG-----GA 2400  
2353 TGCCAGTGCATGAGGAATCCGTTGTATCGAGAGAGCATCTATGAATGAGAGAG 2422  
2401 GACCAGCCCGTGAAGGATGCCGTGTCATCGAGAGAGCTCTCCTGAAATGAGGCCAG 2460  
2423 CAGAGCGCTGTTGACCTTCTTCGCACTTCTCGGTGACAGTGTCTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTCTGACGCTTCTCTACGCTTCTATGTGCCACCGCGGTGCGCATCAGGGG 2520  
2483 CAGTGTCTATACGAGTGCAGGGCATCCCGAGGGCTCCAGGCTTCACACCTGCTCTG 2542  
2521 CAAGTCTAGCTGTCAGTGCAGGGATCCCGAGGGCTCTCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTGTGTTTGGAGACATGAGAAACAAGCTGTTTGTGAGGTGCAGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGACATGAGAAACAAGCTGTTTGGGGGATTCGGCGGAGCGGGCT 2640  
2603 GCTTTTACGTTTGTGATACATTTCTGTTGGTACGCTCTCAGCTTGGACCAAGCAAAAC 2662  
2641 GCTCTCGGTTTGGTGATGATTTCTGTTGGTGACACCTCACTCACTCACTCACTCACT 2700  
2663 CTTCCTCAGCACCTGCTCCATGCGCTTCTGATGATGGGTGCATATAAACTTCAGAA 2722  
2701 CTTCCTCAGGACCTGCTCGAGGTGCTCCCTGATGATGCTGGTGGTAACTTCGGGA 2760  
2723 GACAGTGGTGAACCTTCCCTGTGAGCCTGTGACCTGCTGGGTGTGAGCTCCATACAGCT 2782  
2761 GACAGTGGTGAACCTTCCCTGTGAGACGAGGCTTGGTGGTGGTGGTGGTGGTGGTGG 2820  
2783 GCTGTGCTCACTGCTGCTTCTCCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2842  
2821 GCGGGCCACGGCTATTCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2880  
2843 GTTCTGTGCTACTCAGGTTATGCCAGACCTCAATTAAGACAGCCTCACTTCCAGAG 2902  
2881 GCAGAGCGACTACTCCAGCTATGCCGAGACCTCCATCAGAGCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAGACCATGCGGACAGCTCCTGTGCTGTGCTGTGCTGTGCTGTGCT 2962  
2941 CCGTTCAGGCTGGGAGGACATGCTGCGAACTCTTTGGGGTCTTGGGGTCTTGGGGTCT 3000  
2963 TCAGGCTTATTTCTAGACTTGAGTGAACTGAACTCCAGAGCTGCTGATCAATATATA 3022  
3001 TCAGAGCTGTTTCTGATTTGAGTGAACAGCTCCAGAGCTCCAGAGCTGTCACCACTCA 3060  
3023 CAAGATCTTCTGCTTCAAGCTTCAAGGTTCCATGATGTGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTTCTGCTGAGGCTGACAGGTTTCAAGCATGTGTGCTGCTGCTGCTGCTGCT 3120  
3083 CCAGCTGTTTAGGAAGAACTCACTTCTTCTGGGCACTATCTCCAGCAAGATCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTCTGACAGCGCTCCCT 3180  
3143 CTGCTATGCTATCTCAGAGTCAAGATCCAGGAATGACATAAG-----GCTC 3193  
3181 CTGCTACTCATCTTGAAGAACCAAGAACAGAGGATGCTGCTGGGGGCAAGGGCGCGC 3240  
3194 TGCTCTCTTCTCTCTGAGCGCGCATTTGGCTCTGCTACAGAGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTCTGCTCTCGAGGCTGTCAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGCTGCTATCTGCTCACTAATAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTCAGACCGGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3314 GCTGTGCGGAGCTCCAGAGCGGACATGACATCTTAAAGCTGAGCTGACCCAGC 3373

Db 3361 GCTGAGTGGAGACTCCCGGGAGCAGCGTGACTGCTCCCTGGAGGCGCGAGCCACCCGGC 3420  
Qy 3374 CTTAAGCACAGACTTTTCAGACCAATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTTCAGAGCACTCTGACTGA 3454  
RESULT 9  
US-10-105-963-1  
; Sequence 1, Application US/10105963  
; Publication No. US20030068818A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Denning, Chris  
; APPLICANT: Clark, A. John  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
; FILE REFERENCE: 731/002  
; CURRENT APPLICATION NUMBER: US/10/105,963  
; PRIOR FILING DATE: 2002-03-21  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION:  
US-10-105-963-1

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
Qy 23 CCGCGCTTTGACACAAATGACCCGCGCTCTGCTTCCCGCGGTGCGCTCTGCTGCGG 82  
Db 40 CCGCGCACCCCGCGATGCGCGGCTTCCCGGTGCGAGCCGTGCTTCTTCTGCTGCG 99  
Qy 83 CAGCCGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCTTGGGCGCCGAGGG 142  
Db 100 CAGCCACTACCGAGAGTGTGCGCTGGCCACCTTGTGCGGCGCTGCGGCGCCAGGG 159  
Qy 143 CAGCGGCTTGTGCAACCCGCGGAGCCCGAGATCTACGACATTTGCTTGGCCCATGCTCT 202  
Db 160 CTGGCGGTGTGTCAGCGCGCGGAGCCCGCGGCTTCCGCGCGCTGTGCGCGCTGCTCT 219  
Qy 203 AGTGTGCTGCTGCTGCGGCTCACAGCTTCCAGCTTCCCTTCCACCGAGGTGTC 262  
Db 220 GGTGTGCTGCTGCGGAGCGACGCGCGCCCGCGCGCTTCTTCCGCGAGGTGTC 279  
Qy 263 ATCCCTGAAAGCTGTGTGGCAGGCTTGTGAGAGACTGTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTGTGAAGAGTGTGTGGCCCGAGTGTGTCAGAGCTGTGCGAGCGCGCGCGAGAA 339  
Qy 323 CTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 382  
Db 340 CTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 399  
Qy 383 CACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442  
Db 400 CACCACAGCGTGGCGAGCTACCTGCGCCAAACACCGTACCGACGCGACTGCGGGGAGCGG 459  
Qy 443 TGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502  
Db 460 GCGTGTGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
Qy 503 CTGTGCTCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562

520 CTGCGCGCTCTTTGTGTGGTGGCTCCCACTCGGCTACCAAGTGTGGGCGCCGCGT 579  
553 GTACCAATTTGTGCCACCAACGATATCTGGCCCTCTGTCTCGCTAGTTACAGGCCAC 622  
580 GTACCACTCGCGCTGCCACTCAGGCCCGGCCGCCACACGCTAGTGGACCCGAG 639  
623 CGAGCCCGTGGGAGGAATTTCACTAACCTTAGTTCTTACAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGCGCTTGAACCATAG 672  
683 TCGCCAGGAAGCACCAGAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGATCTGAG 742  
673 CGTCAGGAGGCGCGGGTCCCTTGGCTTGCAGCCCGGGTGCAGAGAGGCGGGG 732  
743 TCTCACAGTACAGTGTGCTTCACTAGAGAGCCAGATGATCTCTGTCTCCAGAGT 802  
733 CAGTGGCCAGCGAAGTCTGCGCTTGGCCAAAGAGCCAGGCTGGGCGCTGCCCTGAGCC 792  
803 GGAGAGGAGCCACACAGGAGGCTACCAACCCCATCAGGCAATCATGAGTGGTGCAG 862  
793 GGAGCGGAGCGCGTGGGAGGGTCTGGGCCACCCCGGCGAGACGCTGGACCGAG 852  
863 TCGTCTCGGTCCCGGAGTGCCTACT-----GCAGAGAAAGATTTGTCTTTAA 913  
853 TGACCGTGGTTCTGTGTGTGTCACTGCGCAGACCCCGGAAAGAGCCACTCTTTGA 912  
914 AGGAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGTGTGTGTAAACAGGCCAG 970  
913 GGTGCGCTCTTGACGCGCCACTCCACCCATCCGTGGGCGCCAGCACCCAGCGGG 972  
971 CTCCACAT---CTGTGTGTCAACACCCCGCCAAATGCTTTTTCAGTCTCAGGCCATTTAT 1027  
973 CCCCCATCCACATCGCGGCCACACAGCTCTCTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGAGATGGCCAGAGGCTTAAACCCCTCATTT 1087  
1033 CGAGACCAAGACATTTCTTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
1088 CTAATCTCAGCAACCTCCAGACCTTAATTTGATGGGGCCAGAGACTGTGTGAGATCATTT 1147  
1090 CTAATCTCAGCTCTGTAGGCCCGAGCTGACTGTGGCTCGAGGCTGTGTGAGACCATTT 1149  
1148 TGTGGCTCAAGCGCTAGGACATCAGACCACTCTGTGAGACACACGCTATTCGGCTG 1207  
1150 TGTGGGTTCCAGGGCCCTGGATSCCAGGAGCTCCCGCAGGTTTGGCGCTGCCCCAGCG 1209  
1208 ATACTGGCAGATGGGCCCTCTTCCACAGCTGCTGTGTGAACCATGACAGGTGCCAATA 1267  
1210 CTACTGGGCAATGCGGCCCTGTCTTGGAGCTGCTTGGGAACACGCGCAGTGGCCCTTA 1269  
1268 TGTGAGATCTCTCAGGTCAATTCAGGTTTCGAACAGCAAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGTCTCTCAAGAGCGACTGCCCGCTCGAGTGTCCGCTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACAGAGCC 1342  
1330 CTGTGCCCGGAGAGCCCGAGGCTGTGGGCGGCCCGGAGGAGGAGACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CGGTGCTGTGTGAGTGTCTCCGACAGCAGCAGCGCCCTGGCAGGTGTACGGCTTGT 1449  
1403 TGGGCTGTCTGTCAAGTGTGTGTCTGTGTCTGTGGGTACAGGCAATATGAGCG 1462  
1450 GGGGCTGTCTGTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1509  
1463 CGGCTTTCTTAAGAACTTAAAGAGTTCATCTGTGTGGGAAATACGGCAAGTATCACT 1522  
1510 CGCTTCTCAGAAACACCAAGAGTTCATCTCCCTGGGAGGAGCATGCCAAGTCTCGT 1569  
1523 GCAGAACTGATGGAGAGTAAAGTAGAGATGTCATCTGGCTCGGAGAGCCCGGG 1582

1570 GCAGGAGCTGACGCTGAAGATGAGCGTGGCGGAGCTGCGCTTGGCTGCCAGAGGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGACAGACACCGCTGTGAGGAGAGGATCTCTGGGTACGTTCT 1642  
1630 GGTGTGCTGTGTTCGCGCGCAGACACCGCTCTGCTGAGGAGATCTCTGGCAAGTCTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGTGTACAGTGTCTAGGTCACTCTTTTACATCAGAGA 1702  
1690 GCACTGGCTGATGAGTGTGTACGTGTGAGCTGTCTCAGGCTCTTTTATGTCAACGA 1749  
1703 GAGCATTTCAGAAAGACAGGCTCTTTCTTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACACGTTTCAAAAGAACAGGCTCTTTTCTTACCGAAGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTGTAGTGTGAGCAACACTGTGAGAGAGTGGGTACCGGAGCTGTACAGAGAGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGTGAGTGGGGAGCTGTGGAAAGCAGA 1869  
1823 GGTGAGCATCACAGGACACCTGTGCTAGCCATGCCCCTCTGCAGACTGCGCTTCATCCC 1882  
1870 GGTGAGCAGCATCGGGAAGCCAGGCCCGCTGTGACGTCACACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGCGCTCGGCGCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCCTGACGCGCTCGGCGCATTTGTGAACATGAGTACGTCGTGGGAGCCAGAACGTT 1989  
1943 GGCAGAGGAAGCAGCGCCAGCATTTTCAACCCAGCGCTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGAGGAAAGAGGCGCGAGGCTCTACCTCGAGGTGAAGGACACTGTTCAGGCTGCT 2049  
2003 CAATATGAGCGCAAAAACATCTCTCACTTATGGGCTTCTCTGTACTGGGTATGAATGA 2062  
2050 CAATACGAGCGCGCGCGCCCTCTCGGCGCGCTCTGTGCTGGCGCTGGAGCA 2109  
2063 CATCTACAGACCTCGCGCGCTTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
2110 TATCCAGGCGCTCGGCGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
2123 GATGTATTTTGTAAAGCAGATGTGACCGGGCCCTATGATGCCATCCCCCAGGGTAAGCT 2182  
2170 GCTGTATTTGTCAAGTGGATGTGACGGCGCGTACGACACCATCCCCCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCACTA 2242  
2230 CAGCGAGGTCAATCCAGCATCATCA---AACCCAGAACAGTACTGCTGCTGCTGCTGCTA 2286  
2243 TGCACTGCTCCGAGAGATAGCCAGGCGCAAGTCCACAAAGTCTCTTATGAGACAGGTCAAC 2302  
2287 TGCGTGTCCAGAGCGCGCCATGGGCACTCGGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCGAGCATACATGGGCGAGTCTCTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGGCAGTGCATCAGAGAACTCCGTTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACAGCGCCCTGAGGAGTGGCGTGTCTATCGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTCTTTGACTTTCTTCTGCACTTCTCTGCGTCAAGTGTGCTGTAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTTCAGAGCTTCTCTAGGCTTCTATGTGCCACACGCGCTGCGCATCAGGGG 2520  
2483 CAGGTGTATACGAGTGCAGGGCATCCCCAGGGCTCCAGGCTATCCACCTCTCTCTG 2542  
2521 CAGTCTCAGTCCAGTGCAGGGATCCCGAGGCTCCATCTCTCTCAGCTCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAACAGTGTGTTGTGAGGTGCGAGCGGATTTGGGTT 2602  
2581 CAGCTGTGCTACGCGACATGAGAAACAGTGTGTTGCGGGAATTCGCGGAGACGGCT 2640  
2603 GCTTTTACGTTTGTGTGACTTCTGTGTGTGAGCGCTCTCTTGGACCAAGCAAAAC 2662  
2641 GCTCCTCGCTTGTGTGATGATTTCTTGTGTGTGACCTCACCTCACCCAGCGGAAAC 2700



323 CGTGTGCTTTTGGCTTTGAGCTTGCTTAACGAGGCCAGAGGGCGGGCTTCCCATGSCCTT 382  
340 CGTGTGCTGCTTGGCTTTGCGCTTCTGCTGAGCGGGCCCGGGGGGGCCCGCCGAGSCCTT 399  
383 CACTAGTASCTGGGTAGCTACTTTGCCCAACACATGTTTATTGAGACCTTGGCTGCTAGTGG 442  
400 CACACACAGCGTGGCAGCTACTTCCCAACACACCGTGAACGACGCACTGGGGGAGCGG 459  
443 TGCATGATGCTACTGTGAGCCGAGTGGGGAGCAGCTCTGCTGCTACTCTGCTGTCACACA 502  
460 GCGGTGGGGGGTGTGCTGCGCCCGCTGGGCGACGACGTGCTGCTTCACTGCTGTCACG 519  
503 CTGTGCTCTTATCTTCTGCTGCCCCCAGCTGCTACAGCTGTGCTGCTGCTGCTGCTGCT 562  
520 CTGCGGCTCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAAAATTTGTGCGCACCGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCAGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639  
623 CGACCCGCTGGCGAGGATTTCACTTAACCTTAGCTTCTTACAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGATGC-----GACCGGCTGGAAACCATAG 672  
683 TGGCAGGAAGCACCGAAACCCCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 742  
673 CGTCAGGAGGCGGGGGTCCCTTGGGCTGCGACGCGCGGCTGCGAGGAGGGCGGGGG 732  
743 TCTACCACTAGTACAGTGTGCTTCACTAGTACAGGCGCAGATGCTATCTGCTCCGAGAT 802  
733 CAGTCCAGCGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
803 GGAGGAGGAGCCCCACAGGAGGAGTGTCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 EGAGGGAGCGCCGTTGGGAGGAGGCTCCTGGGCGCCACCGCGGAGGAGCGGTGGAACGAG 852  
863 TCTGCTCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
914 AGGAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGGTGGCTCTCTGCGACGCGCACTCCACCCATCCGTGGGCGCGCAGCACACCGCGGG 972  
971 CTCACAT---CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCCCCCATCCCATCGCGGCCACACGCTCCCTGGGACACGCTTGTCCCGGCTGTACGG 1032  
1028 TGAGACACAGCATTTCTTACTTCCAGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTTCTACTTCTCAGGCA---CAAGAGAGAGCTGGCGCCCTCTT 1089  
1089 CTTACTAGCAACTCTGAGGCGCCAGCTGACTGGGCTCGGAGGCTGCTGAGACCATCTT 1147  
1090 CTTACTAGCTCTCTGAGGCGCCAGCTGACTGGGCTCGGAGGCTGCTGAGACCATCTT 1149  
1148 TGTGGCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACCAACCGTCTATCGGCTG 1207  
1150 TGTGGTTCCAGGCGCTGATGCTGAGGAGCTGCTGAGGAGTGGCGGCTGCTGCGCCAGCG 1209  
1208 ATACTGGCAGATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
1210 TCTACTGGCAATTCGCGCCCTGTTTCTGAGCTGCTTGGGAAACCAAGCGCAGTGGCCCTA 1269  
1268 TGTGAGATCTCTCAGTCACTGAGGCTTCTGAAACAGCAACCAAGAGTGAAGATGC 1327  
1270 CGGGGTGCTCTCAAGAGCGCACTGCGCGCTGAGCTGCGGTCAACCCAGCAGCGGCTGT 1329  
1328 CT-----TGAAACACGAGCC 1342  
1330 CTGTGCGCGGAGAGCCCGAGGCTCTGTGCGGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAG 1389  
1343 ACCGACCTCATGGATTTGCTCCGCGCTGCAACAGCAGTCCCTGCGAGGATAATGGTTTCT 1402

Db 1390 CCGTGCCTTGGTGCAGCTGCTCGCCAGCACAGAGCCCTTGGCAGGTGTACGGCTTCGT 1449  
QY 1403 TCGGGCGCTGCTCTGCAAGGTGGTGTCTCTAGTCTCTGGGTACACAGGCACAATGAGCG 1462  
Db 1450 GCGGGCTGCTCGCGCGCTGCTGCGCCAGGCTCTGGGCTCCAGGCAACAGAAAG 1509  
QY 1463 CCGCTTCTTTAAGAACTTAAGAACTTATCTCTGTTGGGAAATACGGCAGCTACT 1522  
Db 1510 CCGCTTCTCAGGAAACCAAGAAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
QY 1523 GCAGGAACTCATGTGAAGATGAAGATAGGATGCCACTGGCTCCGCGAGCAGCAGCCGCG 1582  
Db 1570 GCAGGAGCTGACCTGGAAGATGAGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629  
QY 1583 GAAGGACCGTGTCCCGCTGCGAGAGCAGCTGTGAGGAGAGGATCTTGGCTACTTCT 1642  
Db 1630 GGTGGCTGTGTTCCGCGCCGAGCAGCAGCTCTGCTGAGGAGATCTTGGCCAAAGTCT 1689  
QY 1643 GTTCTGCTGATGGAACACATACGTGTACAGCTGTAGTTCATTTTATCATCACA 1702  
Db 1690 GCATGGCTGATGAGTGTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749  
QY 1703 GAGCACATTCAGAAAGAACAGGCTCTTCTTACCGTAAGATGTGTGAGAGCAAGCTGCA 1762  
Db 1750 GACCACTTTCAAAAGAACAGGCTCTTTTCTTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
QY 1763 GAGCATTTGAGTCAGGCAACACCTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
Db 1810 AAGCATTTGAAATCAGACAGCATTTGAAAGAGGTGACGTGCGGAGCTGTGGAAGAGCA 1869  
QY 1823 GGTGAGCATCACAGGACACCTTGGCTAGCCATGCCATCTGCAGACTGCGTTTCATCCC 1882  
Db 1870 GGTGAGGACATCGGGAAGCAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929  
QY 1883 CAGCCCAACGGGCTGCGGCCCATTTGGAACATGATTTAGCATGGGTACAGAGCTTT 1942  
Db 1930 CAGGCTGAGCGGCTGCGGCCGATTTGAAACATGACTACGCTGCTGGGAGCCAGAACGTT 1989  
QY 1943 GGGCAGAAAGAACAGGCGCCAGCATTTACCCAGGCTCTCAAGACTCTCTTCAGCATGCT 2002  
Db 1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACTCGAGGCTGAAGGCACATGTTTCAGCGTGT 2049  
QY 2003 CAATATGACGAGCAAAACATCTCACTTATGAGGTCTTCTGCTGCTGCTGCTGCTGCTGCT 2062  
Db 2050 CAATAGAGCGGCGCGGCCCTTCTGCGGCTCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCT 2109  
QY 2063 CATCTACAGGACCTGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
Db 2110 TATCCAGAGGCTTGGCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
QY 2123 GATGTATTTTAAAGGAGATGTGACCGGGGCTTATGATGCCATCCCGCCAGGTTAAGCT 2182  
Db 2170 GCTGTATTTTCAAGGTGATGTGACGGGCGCTGACACACCATCCCGCCAGGAGCT 2229  
QY 2183 GGTGGAGGTTTGTCAATATATGATCAGGCACTCGAGAGCAGTACTGTATCCGCGAGTA 2242  
Db 2230 CAGGAGGCTCATGCCAGCATCATCA---AACCCAGAACATGCTACTGGTGGCTGCTGCT 2286  
QY 2243 TGCAGTGGTCCGAGAGATAGCCAGGCCAAGTCCAAAGTCTTTTAGAGAGCAGGTCAAC 2302  
Db 2287 TGGGTTGGTCCAGAGCGCGCCCATGGCAGCTGCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
QY 2303 CACCTCTCTGAGCTCCAGCCATACATGGGCGAGTCTTAAAGCATCTGAGGATTCAGA 2362  
Db 2347 TACCTTACAGAGCTTCCAGCGGTACATGCGAGTTCGTGGCTCACCTGCGAG-----GA 2400  
QY 2363 TGGCAGTGCCTGAGGAACTCCCGTTGCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
Db 2401 GACCCAGCCGCTGAGGAGTGGCGTCTGCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
QY 2423 CAGCAGGCTGTTGACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482





143 CAGCGGCTTGTGCAACCCGGGAGCCGAAAGATCTACCGACATTTGGTTGCCAATGCT 202  
160 CTGGCGGCTGTGTGAGCGCGGGAGACCCGGGGGCTTTCCGCGCTGGTGGCCAGTGCT 219  
203 AGTGTGATGACATGGGGCTCAGAGCTCCACCTGCGGACCTTTCTTCCACAGAGTGC 262  
220 GGTGTGCTGCTTGGAGCGCAGCGCGCCCGCGCGCCCTCTTTCGCGAGGTGC 279  
263 ATCCCTGAAGAGTGTGTGCGAGGTTGTGAGAGACTTGTGAGCGCAACAGAGAAA 322  
280 CTGCTCTGAAGAGTGTGTGCGCGAGTGTGAGAGGCTGTGCGAGCGCGCGGAGAA 339  
323 CGTGTGCTTTTGGCTTGTGAGTGTCTTAAGAGGCGAGAGCGGACCTCCATGCGCTT 382  
340 CGTGTGCTTTTGGCTTGTGAGTGTCTTAAGAGGCGGCGCGGCGGCGGCGGCTT 399  
393 CACTAGTAGCGTGTGAGTACTTTGCCCAACACTGTATTTAGAGACCTTCGTGTGAGTGG 442  
400 CACCACAGCGTGTGCGAGTACTTGTCCCAACACCGTGAACGACACTGCGGGGAGCGG 459  
443 TGCATGATGCTACTGTGTGAGCGAGTGGGAGCGAGCTGTGCTACTGCTGTGCGACA 502  
460 GCGGTGGGGCTGTGCTGTGCGCGCGTGTGGGCGAGCGTGTGCTGTGCTGTGCGACG 519  
503 CTGTGCTCTTTATCTTGTGTGCGCGCGAGCTGTGCTTACAGGTGTGTGGGTCTCCCT 562  
520 CTGCGCGCTCTTTGTGTGTGCTGCCAGCTGTGCTTACAGGTGTGTGGGTGTGCGCGCGCT 579  
563 GTACCAAAATTTGTGCCACAGCGATCTGGGCTGTGTGCTCGCTAGTTTACAGGCGCAC 622  
580 GTACAGCTGTGCGCTGTGCTGTGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGAG 639  
623 CCGACCGCTGTGGGAGGAATTTCACTAACCCTTAGTCTTTTACAAAGATCAAGAGCAGTAG 682  
640 GCGTCTGGATGC-----GAGCGGCTGTGAAACCATAG 672  
683 TCGCGAGGAACGCGAAACCCCTGGCTTGTCCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 GGTGAGGAGCGCGGGTGTGCTGTGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGG 732  
743 TCTCAGCATACAGTGTGCTTGTGCTGTGAGAGCGAGATGCTATCTTGTCCGAGGT 802  
733 CAGTGTGCGAGCGAAGTGTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 792  
803 GAGAGAGGAGCG 862  
793 GAGCGGAGCG 852  
863 TCCTGCTCGTCCCGCGAGTGTGCTACT-----GCAGAGAGATTTGTCTCTAA 913  
853 TGACGCTGT 912  
914 AGGAAAGTGTGTGACCTGTGCTCTC---TGGGTGTGTGTGTGTGTGTGTGTGTGTGT 970  
913 GGTGTGCTGT 972  
971 CTCACAT---CTGT 1027  
973 CCGCCCATTCACATCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1032  
1028 TGAGACAGACATTTCTTTTACTCCAGGGAGATGGCCAGAGCGCTATAAACCCCTCAT 1087  
1033 CGAGACCAAGACATTTCTTCTACTCTCAGCGA---CAAGGACAGCTGTGGGCTCTT 1089  
1088 CTTACTCAGAACCTTCAGCTTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147  
1090 CTTACTCAGCTCTCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACACCTGTGTGTGTGTGTGTGTGTGTGTGT 1207  
1150 TCTGGGTTCCAGGCGCTGTGATGCCAGGAGATCTCCCGCAGGTTGCCCGCGCTGCCAGCG 1209

Qy 1208 ATACTGGCAGATGGCGCCCTCTTCCAAACAGCTGTGTGTGAACCATGAGAGTGCCAAATA 1267  
Db 1210 CTACTGGCAAAATCGCGCCCTCTTCTGTGAGTGTGTGTGGAACCAACGCGAGTGCCCTTA 1269  
Qy 1268 TGTGACACTCTCTCAGTCACTTGTGAGGTTTTCGAACAGCAAAACAAAGGTGACAGATGC 1327  
Db 1270 CGGGGTGTCTCTCAGAGCCACTGCGCGCTGGAGCTGGGTCAACCCAGCAGCGGTGT 1329  
Qy 1328 CT-----TGAAACACAGGCC 1342  
Db 1330 CTGTGCGCGGAGAACCCAGGCTCTGTGCGCGCCCGGAGGAGGAGACACAGACCC 1389  
Qy 1343 ACCGACCTCATGGAATTTGCTCGGCTGCACAGAGTCCCTGGGAGGTATATGTTTCT 1402  
Db 1390 CCGTGTGTGTGTGAGTGTCTCGCGAGCAGAGAGCCCTTGGGAGGTGTGAGGTGTGCTTGT 1449  
Qy 1403 TCGGGCTGTCTCTGCAAGT 1462  
Db 1450 GCGGGCTGTGCTGTGCGCGGCTGTGTGCGCGCCCGGAGGCTCTGGGCTCCAGGCAACAGAACG 1509  
Qy 1463 CCGCTTCTTTAAGNACTTAAAGATTTTCTCTGTGTGGGAAATACGCAAGCTTATCACT 1522  
Db 1510 CCGCTTCTCTCAGAGACACCAAGAGTTTCTCTCTGTGGGAGGATGCCAAGCTCTGCT 1569  
Qy 1523 GCAGGAACCTGATGTGAGATGAAGATGAGGATTTGCCACTGGCTCCGAGCAGCCCGGG 1582  
Db 1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCAGGAGCCAGG 1629  
Qy 1583 GAAGGACGCTGTCCCGCTGCAGAGCAGCGTGTGAGGAGAGGATCTCTGGTACGTTCT 1642  
Db 1630 GGTGTGTGTGTGTGCGCGCGGAGAGCAGCTGTGCTGTGAGGAGATCTCTGGCAAGTTCT 1689  
Qy 1643 GTTCTGCTGTGTGGAACACATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1702  
Db 1690 GCACTGT 1749  
Qy 1703 GAGCATTTCGAAGAAACAGGCTCTTCTTCTAACCCTAAGAGTGTGTGTGTGTGTGTGT 1762  
Db 1750 GACACGTTTCAAAAGAACAGGCTCTTTTCTTACCGAAGAGTGTGTGTGTGTGTGTGT 1809  
Qy 1763 GACATTGTGAGTCAGCAACACTTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1822  
Db 1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1869  
Qy 1823 GGTGAGGATCAACAGGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1882  
Db 1870 GGTGAGGATCAACAGGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1929  
Qy 1883 CAAGCCCAACGCGCTGTGCGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1942  
Db 1930 CAAGCTGTGCGGCTGTGCGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1989  
Qy 1943 GGGCAGAAGAAAGAGCGCCAGCTTTTCCAGCGCTTCAAGACTCTTTCAGACTGCT 2002  
Db 1990 CCGCAGAGAAAGAGCGCGAGCTCTCACTTCGAGGTGTGAAGCACTGTTCAGCGTGT 2049  
Qy 2003 CACTATGAGCGGACAAACATCTCACTTATGGGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2062  
Db 2050 CAACTAGAGCGGCGCGCGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2109  
Qy 2063 CATCTACAGCATCTGCGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
Db 2110 TATTCACAGGCTGTGCGCACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
Qy 2123 GATGTATTTGTTAAGCAGATGTGACCGGGGCTTATGATGCCATTCGCCAGGTGAGCT 2182  
Db 2170 GCTGTATTTGTCAAGGTGTGTGTGACCGGCGCGTACGACCATTCGCCAGGACAGCT 2229  
Qy 2183 GGTGAGGT 2242  
Db 2230 CAGGAGGT 2286  
Qy 2243 TGCGT 2302

2287 TCCGCTGTCAGAGGCGGCCCATGCGCAGTCCGAGGCCCTTCAGAGCCACAGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAGCATCTGAGGATTGAGA 2362  
2347 TACCTTGAAGAGACCTCCAGCCGCTACATCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGCACTGAGGAACTCGTGTCTATCGACAGAGCATCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGGATGCGGTGCTCATGAGCAGAGCTCCCTGTAATGAGGCCAG 2460  
2423 CAGAGCCCTGTTGACTTCTTCTCCTGACATTCCTCGTCCACAGTGTCTGAAGATTGTTGA 2482  
2461 CAGTGGCTCTTCGAGCTTCTCTACGCTTCATGTGCCACCAAGCGGTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGACGTCAGGACATCCCCAGGCTCCAGGCTCCAGCTATCCACCTGCTCTG 2542  
2521 CAGGTCTACGTCCTGAGTCCAGGGATCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGGAGACATGGAGAACAAAGCTGTTGCTGAGGTGAGGGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGACATGGAGAACAAAGCTGTTGCGGGATTCGGCGGACGGGCT 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGTTGGTACGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCCTGCGTTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCGCAAAAC 2700  
2663 CTTCTCTCAGCACCTTGGTCCATGGGCTTCTGAGTATGGGTGCATGATAACTTGCAGAA 2722  
2701 CTTCTCTCAGGACCTTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGTGACTTGGGAA 2760  
2723 GACAGTGTGAACTTCCCTGTGGAGCCCTGGTACCTCGTGGTGGTGCAGTCCATACAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTAGAGACAGAGGCCCTGGGTGGCACGGCTTTTGTTCAGAT 2820  
2783 GCTGCTCACTGCTGTTTCCCTGTGTGGTGTGCTGCTGAGACTCAGACTTTGAGGT 2842  
2821 GCGGCCACAGGCTTATCCCTGTGTGGGCTGCTGCTGTGATACCGGACCTTGGAGGT 2880  
2843 GTCTGTGACTACTCAGGTATGCCCCAGACCTCAATTAAGACGAGCTCACCTTCCAGAG 2902  
2881 GCAGAGCGACTACTCAGCTATGCCCCGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGCTTCAGGCTGGAGACCAACGAGAAACAGCTCTGCTGCTGCTGCTGGTGGTGAAGTG 2962  
2941 CGCTTTCAGGCTGGAGGAAACATGCTGCAAACTCTTGGGCTCTGCGGCTGGAAGTG 3000  
2963 TCAGGCTCTATTCTAGACTTCAGGTGAAGAGCTCCAGAGCTCTGCATCAATATATA 3022  
3001 TCACAGCTGTTTCTGGATTTGACAGTGAACAGCTCCAGACGCTGTGCACCAACATCTA 3060  
3023 CAGATCTCTGCTTACGGCTCAGGTTCAATGATGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTCTGCTGCTGAGGCTACAGTTCATGCGATGTGTGCTGAGCTCCCATTTCA 3120  
3083 CCAGCGGTGTAGAGAACTCACAATCTTCTGCGCATCATCTCCAGCAAGCATCCTG 3142  
3121 TCAGCAGTGTGAGAGAACCCACATTTTCTGCGGTCTATCTGACAGGCTCCCT 3180  
3143 CTGCTATGCTATCTGAGGTCAAGATTCAGGAATGACATAAG-----GCCTC 3193  
3181 CTGCTACTCCATCTGAAAGCCAGAGCGAGGATGCTGCGGGGCAAGGGCGCGCG 3240  
3194 TGGCTCCTTCTCTGAGCGCGCATTTGGCTCTGCTPACAGGCTTCTGCTCAGCT 3253  
3241 CGGCCCTCTGCCCTCGAGCGCGTCACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGTGCTCATCTGCTCATCTAATAATGTCTCTGCGACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTGACACCGCTGTCACCTAGTGCCTACTCTGCGGTCACTCAGGACAGCCAGCGCA 3360  
3314 GCTGTGCGGAGCTCCAGAGCGCAATGACCATCTTAAAGCTGAGTGAACCCAGC 3373

Db 3361 CTGAGTCGGAGACTCCGGGACGAGCTGACTGCTCGCTGAGGCGCAGCCACCCGCG 3420  
QY 3374 CTAAGCACAGACTTTTCAGACCAATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454  
RESULT 12  
US-10-385-882-1  
; Sequence 1, Application US/10385882  
; Publication No. US2003032409A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA  
; APPLICANT: FARRIS, James  
; APPLICANT: FOSTER, Douglas  
; APPLICANT: O'GRADY, Scott  
; TITLE OF INVENTION: IMMORTAL PORCINE CELLS  
; FILE REFERENCE: 110.01700101  
; CURRENT APPLICATION NUMBER: US/10/385,882  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR FILING DATE: 2002-03-11  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1  
; LENGTH: 4027  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence encoding telomerase reverse transcriptase  
US-10-385-882-1  
Query Match 45.8%; Score 1600.8; DB 12; Length 4027;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCGCGCTTGAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCTGCG 82  
Db 41 CCGCGCACCCCGCGATGCGCGCGCTCCCGCTCCGAGCGCTGCTCTGCTGCG 100  
QY 83 CAGCGCATACCGGAGGTGTGGCGCTGGCAACTTTGTGCGCGCTCGGGCCCGAGG 142  
Db 101 CAGCACTACCGCAGGTGTCTGCGCTGGCACTGTTGTGCGGCGCTGGGCCCGAGG 160  
QY 143 CAGCGGCTTGTGCAACCCGGGACCCGAAAGTCTACCGCACTTTGTGTTGCCAATGCT 202  
Db 161 CTGCGGCTGTGTGAGCGCGGGACCCGGCGGCTTTCCGCGCTGTGCGCCAGTGCCT 220  
QY 203 AGTGTGATGACATGGGCTCAGACCTCCACCTGCGGACCTTTCTTCCACAGGTGTC 262  
Db 221 GGTGTGCTGCTTGGACGACGCGCGCCCGCCCGCGCGCTCTTTCGCGAGGTGTC 280  
QY 263 ATCCTGAAAGAGCTGTGCGCAGGTTGTGCAAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 281 CTGCTGAAGAGCTGTGCGCGCCAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAGAA 340  
QY 323 GGTGTGCTTTTGGCTTTGAGCTGTTAAGAGGCGAGGCGGGGCTCCCATGCGCTT 382  
Db 341 GGTGTGCGCTTTCGGCTTTCGGCTGCTGAGCGGGGCGCGGGGCGCGCGGCGCTT 400  
QY 383 CACTAGTAGCTGCTGAGTACTTTGCCCAACACTGTTATTGAGACCTCTGCTGTCAAGTG 442  
Db 401 CACCACAGCGTGCAGCTACCTGCCCAACACCGTGACCGGACCTGCGGGGAGCGG 460  
QY 443 TGCATGATGCTACTTGTGAGCGAGTGGCGGACGACCTGCTGCTACTGCTGCGACA 502  
Db 461 GCGGTGGGGGCTGCTGTTGCGCGCGTGGCGGACGAGTGTGCTGCTGCTGCGACG 520  
QY 503 CTGTGCTCTTTATCTTCTGCTGCGCGCGCTGCTGCTACAGGTGTGTTGGGTCTCCCT 562  
Db 521 CTGCGGCTCTTTTGTGCTGCTGCTGCTCCAGCTGCGCTACAGGTGTGCGGCGCGCT 580  
QY 563 GTACCAAAATTTGTGCCACACCGGATATCTGCGCCCTCTGTGCTGCTAGTTACAGGCCAC 622

[illegible]

2723 CACAGTGGTGAACCTCCCTGTGGAGCCTGGTACCTGGTGGTGCAGCTCCATCAGACT 2782  
|||||  
2762 GACAGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGTGGCACGGCTTTTGTTCAGAT 2821  
|||||  
2783 GCCTGCTCACTGCCTGTTTCCCTGGTGTGGCTTGCTGCTGGACACTCAGACTTTGGAGGT 2842  
2822 GCCGGCCACGGCTATTTCCTCCCTGGTGGCGCTGCTGCTGGATACCCGACCCCTGGAGGT 2881  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACTCAATTAAGACGAGCCTCAGCTTCCAGAG 2902  
2882 GCAGAGCGACTACTCCAGCTATGCCCGAACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2941  
2903 TGTCTTCAAAGCTGGAGAGACCATGCGGAACAAGCTCCTGTGCGTCTTTGGGTTGAAGTG 2962  
2942 CGGCTTCAAGGCTGGAGGACATGCGTGCGAACCTCTTTGGGCTTTGGGCTGAAGTG 3001  
2963 TCACGGTCTATTCTTAGACTTTGCAAGGTGAACAGAGCCTCCAGACACTGTGCATCAATATATA 3022  
3002 TCACAGCCTGTTTCTGATTTGCAAGTTGAACAGAGCCTCCAGACGGTGTGCCAACATCTA 3061  
3023 CAAGATCTTCTGCTTCAAGCCTACAGGTTCCATGCATGTGTGATTCAGCTTCCCTTTGA 3082  
3062 CAAGATCTCTGCTGTCAGGCGTACAGGTTTTCAGGCATGTGTGCTGCAGCTCCCATTTCA 3121  
3083 CCAGCGTGTTAGGAAGAACCTTCACATTCTTTCTGGGCATCATCTCCAGGCCAAGCATCCTG 3142  
3122 TCAGCAAGTTTGGAGAACCCCAATTTTCTTGGCGGTCACTCTGACACGGGCTCCCT 3181  
3143 CTGCTATGCTATCTCTGAAGTTCAGAAATCCAGGAATGACACTAAAG-----GCCTC 3193  
3182 CTGCTACTCCATCTCTGAAGCAAGAACGAGGAGTGTGCTGGGGGCCAAGGGCGCGC 3241  
3194 TGGCTCTCTTCTCTGGAAGCGCACATTGGCTCTGCTACAGAGCCTTCTGCTCAAGCT 3253  
3242 CGGCGCTCTGCGCTCCGAGGCGGTGTCAGTGGTGTGTCACCAAGCATTTCTGCTCAAGCT 3301  
3254 GGCTGCTCAATCTGTGATCTACAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
3302 GACTCGACACCGGTGTCACTACGTGCGCACTCCTGGGCTCACTCAGGACAGCCACAGCGCA 3361  
3314 GCTGTGCGGAGCTCCAGAGGCGCAATGACCATTCTTAAGCTCAGCTGACCCAGC 3373  
3362 GCTGAGTCGGAAGTCTCCGGGAGCGAGCTGACTGCGCTGAGGCGCGCAACCCGCG 3421  
3374 CCTAAGCACAGACTTTCAGACCATTTTGGACTAA 3407  
3422 ACTGCCCTCAGACTTCAAGACCATCTCTGACTGA 3455

SEQUENCE 32, Application US/09749728B  
Patent No. US2020142457A1  
GENERAL INFORMATION:  
APPLICANT: Umezawa, Akihiro  
APPLICANT: Hata, Jun-ichi  
APPLICANT: Fukuda, Keiichi  
APPLICANT: Ogawa, Satoshi  
APPLICANT: Sakurada, Kazuhiro  
APPLICANT: Gojo, Satoshi  
APPLICANT: Yamada, Yoji  
TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMY  
FILE REFERENCE: 00766.000043  
CURRENT APPLICATION NUMBER: US/09/749,728B  
CURRENT FILING DATE: 2001-09-17  
PRIOR APPLICATION NUMBER: H11-372826  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: PCT-JP00-01148  
PRIOR FILING DATE: 2000-02-28  
PRIOR APPLICATION NUMBER: PCT-JP00-07741  
PRIOR FILING DATE: 2000-11-02  
NUMBER OF SEQ ID NOS: 80

SULT 13  
-09-749-728B-32  
Sequence 32, Application US/09749728B  
Patent No. US20020142457A1  
GENERAL INFORMATION:  
APPLICANT: Umezawa, Akihiro  
APPLICANT: Hata, Jun-Ichi  
APPLICANT: Fukuda, Keiichi  
APPLICANT: Ogawa, Satoshi  
APPLICANT: Sakurada, Kazuhiro  
APPLICANT: Gojo, Satoshi  
APPLICANT: Yamada, Yoji  
TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMYOCYTES  
FILE REFERENCE: 00766.000043  
CURRENT APPLICATION NUMBER: US/09/749,728B  
CURRENT FILING DATE: 2001-09-17  
PRIOR APPLICATION NUMBER: H11-372826  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: PCT-JP00-011148  
PRIOR FILING DATE: 2000-02-28  
PRIOR APPLICATION NUMBER: PCT-JP00-07741  
PRIOR FILING DATE: 2000-11-02  
NUMBER OF SEQ ID NOS: 80



3022 CAGGCGTACAGGTTTCAGCATGTGTGCTGAGCTCCCAATTCATCAGCAAGTTTGAAG 3081  
3099 AACCTCACATTTCTTCTGGGCAATCATCTCCAGCCAGCAATCTCTGCTATGCTATCTG 3158  
3082 AACCCACATTTTCTGCGCTCATCTCTGACAGCGCTCCCTCTGCTACTCATCTG 3141  
3159 AAGGTCAAGATCCAGGAATGACACTAAAG-----GCTCTGGCTCCTTTCTCCT 3209  
3142 AAAGCCAAAGACGAGGATGTCTGCTGGGGCCAGGCGCGCGGCTCTGCGCTCC 3201  
3210 GAAGCCGCAATGCTGTGTACACAGCCCTTCTGCTGCTCAAGCTGGTGTCTCATCTGTC 3269  
3202 GAGGCGGTGAGTGTGTGTCACCAAGCAATCTCTGCTCAAGCTGACTCGACACCGGTGTC 3261  
3270 ATCTCAAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAATCTGCTGTGCGGAGCTC 3329  
3262 ACTAGTGCCTACTCTGCGGTGCTCTAGGAGAGCCCAAGCGCAGCTGAGTGTGGAAGCTC 3321  
3330 CCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGACCCAGCCCTAAGCAGACAGCTTT 3389  
3322 CCGGGAGACGCTGACTGCTGCGCTGAGGCGCGCAGCCAAACCGGCACTGCCCTCAGACTTC 3381  
3390 CAGACCAATTTGGAC 3404  
3382 AAGACCATCTCGGAC 3396

SUBT 14  
-10-205-629-1  
Sequence 1, Application US/10205629  
Publication No. US20030049236A1  
GENERAL INFORMATION:  
APPLICANT: Kassen, Thomas  
APPLICANT: Jensen, Thomas  
APPLICANT: Ratten, Suresh  
TITLE OF INVENTION: Immortalized Stem Cells  
FILE REFERENCE: 006148.00002  
CURRENT APPLICATION NUMBER: US/10/205,629  
PRIOR FILING DATE: 2002-07-26  
PRIOR APPLICATION NUMBER: 60/315939  
PRIOR FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: PA 2001 01148  
PRIOR FILING DATE: 2001-07-27  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 3453  
TYPE: DNA  
ORGANISM: Homo sapiens  
-10-205-629-1

Query Match 45.6%; Score 1593.6; DB 15; Length 3453;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2370; Conservative 0; Mismatches 964; Indels 108; Gaps 9;  
35 CAAATGACCCCGCTCTCTGTCGCCGCGTGGCTCTCTGTCGGCAGCGATACCG 94  
6 CACCAATGCCGCGCTCCCGCTGCCGAGCGTGGCTCTCTGTCGGCAGCGATACCG 65  
95 GAAGTGTGGCGCTGGCAACTTTGTGCGCGCTGCGGCGCCGAGGCGCAGCGGCTGT 154  
66 CGAGTGTGCGCTGGCCAGCTTGTGTCGGCGCTGCGGCGCCGAGGCTGGCGGCTGT 125  
155 GCAACCCGGGACCCGAGATCTACCGACTTTGGTGGCCCAATGCTAGTGTGATGCA 214  
126 GAGGCGGCGGACCCGCGGCTTTCCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTG 185  
215 CTGGGCTCAGAGCTCCAGCTGCGGACTTTCTTCCACAGGTGTCTCTGCTGCTGCTG 274  
186 CTGGGACGACGCGCGCGCGCGCGCTCTCTTCTGCGCAGGTGCTGCTGCTGCTGCTG 245  
275 GCTGTGGCGCGGCTGTGTGAGAGACTCTGCGAGCGCAACGAGAGAAACGCTGCTGCTT 334

Db 246 GCTGTGCCCCGAGTGTGTCAGAGCTGTGCGAGCGCGCGCGAAGACGCTGTGCGCTT 305  
QY 335 TGGCTTTGAGCTGCTTAACGAGGCGAGAGCGGCGCTCCCATGCGCTTCTCACTAGTAGCGT 394  
Db 306 CGGCTTGCCTGCTGTGACGCGGCGCGCGCGCGCGCGCGCTTCAACACCGAGCT 365  
QY 335 CGGTAGCTACTTGGCCCAACACTGTTATTGAGACCTGCGTGTGCTGCTGCTGCTGCTGCT 454  
Db 366 GCGGAGCTACTCTGCGCCCAACACGCTGACCGACGCTGCGGCGGAGCGGCGGCTGCGGCT 425  
QY 455 ACTGTTGAGCGAGTGGGCGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514  
Db 426 GCTGCTGCGCGCTGCGGCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485  
QY 515 TCTTCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574  
Db 486 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545  
QY 575 TGCCACACGAGTATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634  
Db 546 GCTGCGCACTCAGGCG 605  
QY 635 CAGGAATTTCTACTAACTTAGTCTTCAACAGATCAAGAGCAGTAGTGTGCGCAGGAAC 694  
Db 606 C-----GAAACGCGCTGGAACCATAGCTGAGGAGGCG 638  
QY 695 ACCGAACCCCTGSCCTTGCCATCTCGAGTACAAAGAGGCATCTGAGTCTCACCAGTAC 754  
Db 639 CGGGTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698  
QY 755 AAGTGTGCTTCACTAAGAGCGCAGATGCTATCTGCTGCGGAGAGTGGAGAGGAGGAC 814  
Db 699 AAGTGTGCTTGGCG 758  
QY 815 CCAGGAGGTGCTACCAACCCATCAGGCAATCATGCGTGCAGGCTGCTGCTGCTGCTGCT 874  
Db 759 GCTTGGCGAGGCTCTGCG 818  
QY 875 CCGGAGGTGCTTACT-----GCAGAGAAAGATTTGCTCTCTTAAAGGAAAGGTGTC 925  
Db 819 CTGTGTGCTGCTCACTGCGCAGACCGCGCGGAGAGCCACTCTTTGGAGGCTGCGCTCTC 878  
QY 926 TGACTGAGTCTCTC-----TGGGTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979  
Db 879 TGGCAGCGCGCACTCCCAACCATCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 938  
QY 980 TCTGCTGTCAACCCCGCGCAAAATGCTTTCAGCTCAGGCCATTTATTGAGACGAGACA 1039  
Db 939 ATCGCGGCGCACGCTCCCTGCGGACAGCGCTTCTCCCGGCTGACCGCGAGACCAAGCA 998  
QY 1040 TTTCTTTTACTCCAGGAGATGCGCAAGAGCGCTTAAACCCCTCTATTCTTCTACTCAGCAA 1099  
Db 999 CTTCTCTTCTCTCTCAGGCGA---CAAGGAGCAGCTGCGCGCTCTCTTCTACTCAGCTC 1055  
QY 1100 CTTCCAGCTTAACTTGAATGGGCGCAGGAGCTGCTGCGGAGATCATCTTTCTGGGCTCAAG 1159  
Db 1056 TCTGAGCGCGCAGCTGACTGCTGCGGCTGCGGAGTCTGCGGAGACCATCTTTCTGGGTTCCAG 1115  
QY 1160 GCCTAGGACATCAGGACCACTCTGCGAGGACACACCGCTCTATGCGGTCGATAGTGGCAGAT 1219  
Db 1116 GCGCTGATGCCAGGACTCCCGCAGGTTGCCCGCTGCGCGCGCTACTGGCAAT 1175  
QY 1220 GCGGCGCTGTTTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279  
Db 1176 GCGGCGCTGCTTCTGAGCTGCTTGGGAAACACGCGGAGTGGCGCTTACGGGCTGCTCT 1235  
QY 1280 CAGTGTCACTTTCAGGTTTTCGAAACAGCAACCAACAGGTGACAGATGCT----- 1329  
Db 1236 CAGACGCACTGCGCGCTGCGAGCTGCGGTACCCCGCAGCGAGCTGCTGCTGCTGCTGCTGCT 1295  
QY 1330 -----TGAACACAGCGCGCAACGAGAGAAACGCTGCTGCTT 1354



1296 GAAGCCCCAGGGCTCTGTGGGGCCCCGAGGAGGAGACACAGACCCCGCTGCCTGGT 1355  
1355 GGAATTTGCTCCGCTGCACAGCAGTCCCTGGGAGGTATATGGTTTTCTTCGGGCTGTCT 1414  
1356 GCAGCTGCTCCGCGCAGACACAGCAGCCCTGGCAGGTGTACGGCTTCGTGGGGCTGCCT 1415  
1415 CTGCAAGGTGGTCTCTGCTAGTCTCTGGGGTACCAAGGACACATGAGCGCCCTCTTTAA 1474  
1416 GCGCCGCTGGTCCCGCCCGAGGCTCTGGGGCTCCAGGGACACAGAACCCGCTTCCTCAG 1475  
1475 GAACTTAAAGAACTTCATCTCGTTGGGAAATACCGCAAGCTATCACTGCAGGAACTGAT 1534  
1476 GAACACCAAGAGTTTCACTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGAC 1535  
1535 GTGGAAGATGAAGTGAAGATGACCATGCTGGCTCCGACAGCCCGGGGAGGACCGTGT 1594  
1536 GTGGAAGATGAGGCTCGGGAGCTGCGCTTGGCTGGCGAGGAGCCGAGGGTGTGGCTGTGT 1595  
1595 CCGCGCTGCAGAGACCGTCTGAGGGAGAGGATCCTGGCTAGCTTCTCTGTCTGGCTGAT 1654  
1596 TCGGCGCGAGAGCAAGCTCTGGTGAGGAGATCTGGCCAGATTCCTGGCTGAT 1655  
1655 GGAACATAGCTGGTACAGCTGTAGGTGATCTCTTTTACATCAAGAGAGACATTTCCA 1714  
1656 GAGTGTGTACGCTCGAGCTGTCTAGGTCTTTCTTTTATGTACGGAGACACGTTTCA 1715  
1715 GAAGAACAGGCTCTCTCTACCGTAAAGTGTGGAGCAAGCTGCAGAGCATTTGAGT 1774  
1716 AAGAACAGGCTCTTTTCTACCGCGGAGTGTCTGGAGCAAGTTGCAAGCATTTGAGT 1775  
1775 CAGCACAACCTTTGAGAGATGCGGCTACCGGAGTGTCAAGAGAGGAGTCAAGCATCA 1834  
1776 CAGACAGACTTGAAGAGGTGCAGCTCGGGAGTGTGGAAGCAGAGGTCAAGCAGCA 1835  
1835 CCAGGACACTGGCTAGCATGCCATCTGCAGCTGCGCTTCACTCCCGAGCCCAAGCG 1894  
1836 TCAGGAGCAGGCGCGCTCTGTGAGTTCAGCTCCAGCTTCCCTTCACTCCCGAGCTGAG 1895  
1895 CTTGCGGCGCATTTGTGAATGAGTTATAGCATGGGTACAGAGCTTTGGGAGAGGAA 1954  
1896 GCTGCGGCGGATTTGGAACATGAGTACGTCTGGGGAGCCAGAACGTTCCGAGAGAAA 1955  
1955 GCAGGCGGAGATTTACCGAGGCTCTCAGACTCTCTTTCAGCATGCTCACTATGAGCG 2014  
1956 GAGGCGCGAGCGTCTACCTCGAGGCTGAGGCACTGTTCAAGCGTCTCACTACGAGCG 2015  
2015 GACAAAACATCTCACCTATGGGCTCTCTGATCGGTATGATGATCATCTACAGGAC 2074  
2016 GCGCGGCGCGCGCTCTCTGGGGCTCTGTGCTGGGCTGGAGCATATCCACAGGCG 2075  
2075 CTGCGGCGCTTTGT 2134  
2076 CTGCGGCACTTCTGT 2135  
2135 TAAGCGAGATGTACCGGGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 2194  
2136 CAGGTGATGTGACGGCGGTACGACACATCTCCCGAGGACAGGCTCACGGAGTCA 2195  
2195 TGCCAAATATGATAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTATGAGTGGTGG 2254  
2196 CGCCAGCATCATCAACCCCC---AGAACAGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2252  
2255 GAGAGATAGCAAGCCCAAGTCCACAGTCTTTAGGAGCAGGTCACACCCCTCTCTGA 2314  
2253 GAAGCGCCCATCGGCGAGCTGCGCAAGGCTTCAAGAGCAGCTCTTACCTTGAAGA 2312  
2315 CTTCCAGCATATGATGGGCGAGTTCTTTAAGCATCTGCAGATTCAGATGCCAGTGCAT 2374  
2313 CTTCCAGCGTACATGCAAGTTCTGGTGTACCTGAG-----GAGACCAGCCCGCT 2366  
2375 GAGGAATCTCGTTGTATCGAGGAGAGCATCTCTAATGATGAGAGCAGCAGCGCTGTT 2434  
2367 GAGGATGCGCTGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCGCAGGTGGCTCTTT 2426

RESULT 15

US-10-105-616-1

; Sequence 1, Application US/10105616

2435 TGACTTCTTCTGCACTTCTCTGCTGCAAGTGTCTGTAAGATTGGTGACAGGTGCTATAC 2494  
2427 CGACGCTTCTTCTGCACTTCTCTGCTGCAAGTGTCTGTAAGATTGGTGACAGGTGCTATAC 2486  
2495 GCAAGTCCAGAGGCAATCCCGCAGGGCTCAGCCTATCCACCGCTCTCTGCAAGTGTGTTT 2554  
2487 CAGTGTCCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTCTCTGCAAGTGTGCTA 2546  
2555 CGGACATGAGAAACAGCTGTTTGTGAGGTCCAGCGGATGGGATGGGTTGCTTTTACGTTT 2614  
2547 CGGCGACATGAGAAACAGCTGTTTGGGGGATTCGGCGGAGCGGCTGCTCTCTGCGTTT 2606  
2615 TCTCATGACTTCTTGTGTGAGCGCTCACTTGGACCAAGCAAAAACCTTCTCTCAGCAC 2674  
2607 GGTGATGATTTCTTGTGTGACACCTCACTCCACCGCAAAAACCTTCTCTCAGCAC 2666  
2675 CTTGTCATGCGCTTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2734  
2667 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2726  
2735 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2794  
2727 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2786  
2795 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2854  
2787 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2846  
2855 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2914  
2847 CTTGTCGAGGTCTCTGAGTATGGGTGATATAAATTGCAAGAACAGTGTGTA 2906  
2915 TGGGAGAGCATTCGCGAAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2974  
2907 TGGGAGAGCATTCGCGAAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2966  
2975 TCTAGACTTGCAGGTGAACAGGCTCCAGAGTCTGATCAATATATACAGATCTTCTT 3034  
2967 TCTGAGTTTGCAGGTGAACAGGCTCCAGAGTCTGATCAATATATACAGATCTTCTT 3026  
3035 GCTTCAAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3094  
3027 GCTGAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3086  
3095 GAAGAACCTTCAATTTCTGCGGATCATCTCAGGCAAGCATCTGCTGCTGCTGCTGCT 3154  
3087 GAAGAACCTTCAATTTCTGCGGATCATCTCAGGCAAGCATCTGCTGCTGCTGCTGCT 3146  
3155 CTTGAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3205  
3147 CTTGAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3206  
3206 TCTGAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3265  
3207 CTTGAGGCTACAGTTTCCATGATGTGATTCAGTTCCTTCCCTTCCCTTCCCTTCCCT 3266  
3266 TGTCTCTCAATATCTCTGCGGCTCTGAGGACAGCCCAAAATCTGCTGCTGCTGCTGCT 3325  
3267 TGTCTCTCAATATCTCTGCGGCTCTGAGGACAGCCCAAGCATCTGCTGCTGCTGCTGCT 3326  
3326 GCTTCCAGAGGCGACATGACCATCTTAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3385  
3327 GCTTCCGAGGCGACGCTGATGCTTCCCTGAGGCGCGAGCAACCCGCGCTGCTGCTGCT 3386  
3386 CTTTTCAGCATTTTGGACTAA 3407  
3387 CTTCAAGACCATCTCTGGACTGA 3408

Publication No. US20030175967A1  
GENERAL INFORMATION:  
APPLICANT: Genon Corporation  
INVENTOR: Clark, A. J.  
APPLICANT: Denning, Chris  
APPLICANT: Cui, Wei  
APPLICANT: Zhao, Debbiao  
TITLE OF INVENTION: Vectors for Telomerizing Nuclear Donor Cells and Improving the Efficiency of Nuclear Transfer  
FILE REFERENCE: 732/002  
CURRENT APPLICATION NUMBER: US/10/105,616  
CURRENT FILING DATE: 2003-01-31  
PRIOR APPLICATION NUMBER: US Provisional Application 60/277,749  
PRIOR FILING DATE: 2001-03-21  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 13766  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificial construct comprising human TERT, myeloproliferative sarcoma virus (MPSV) promoter, and vector components  
-10-105-616-1

Query Match 45.6%; Score 1593.6; DB 13; Length 13766;  
Best Local Similarity 58.9%; Pred. No. 0;  
Matches 2370; Conservative 0; Mismatches 964; Indels 108; Gaps 9;

35 CACATGACCCCGGCTCTCTGCTGCGCCCGGCTGCGCTCTCTGCTGCGCAGCGATACCG 94  
3985 CACCATGCGCGCGCTCCCGCTGCTGCGCAGCGCTGCGCTCTCTGCTGCGCAGCGATACCG 4044

95 GGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCTGCGCGCGCTGCGCGCGCTGCTGT 154  
4045 CGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCTGCGCGCGCTGCTGTGCGCGCTGT 4104

155 GCACCCCGGGACCCGAGATCTACCGACATTTGTTGCGCGCAATGCTGTGCGCATGCA 214  
4105 GCAGCGCGGGACCCGCGGCTTTCCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGTGC 4164

215 CTGGGCTCAGAGCTCCACCTGCGGCTCTTCTTCCACAGGCTGTCCTGCGGAGGA 274  
4165 CTGGGACGACG 4224

275 GCTGTGCGCGAGGTTGTGCGAGACTCTGCGAGCGCAACGAGAGAAACGCTGTGCGCTT 334  
4225 GCTGTGCGCGAGTGTGCGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 4284

335 TGCGTTGAGCTGCTTAACGAGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 394  
4285 CGGCTTGGCGCTGTGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4344

395 GCGTAGCTACTTGGCCCAACACTGTTTATGAGACCGCTGCTGCTGAGTGGTGCATGGATGCT 454  
4345 GCGCAGCTACTTGGCCCAACACTGTTTATGAGACCGCTGCTGCTGAGTGGTGCATGGATGCT 4404

455 ACTGTGAGCGAGTGGGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514  
4405 GCTGTGCGCGCGCTGCGGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4464

515 TCTTCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574  
4465 TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4524

575 TGCCACCGAGGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634  
4525 CGCTGCCACTCAGGCG 4584

635 CAGGAATTTCATACCTAGTGTCTTACACAGATCAAGAGCAGTGTGCTGCGCAGGAGC 694  
4585 C-----GAAACGGGCTTGGAAACATAGCTGCGAGGAGC 4617

QY 695 ACCGAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACAGTAC 754  
Db 4618 CGGGGTCCCCCTGGGCTTGCAGCCCGGGTGGAGAGCGGGGCGAGTCCACCGC 4677  
QY 755 AAGTGTGCTTCAAGTAAGAGGCGAGATGCTATCTCTCCCGAGAGTGGAGGAGGACC 814  
Db 4678 AAGTGTGCTTCCCGAGAGGCGCGCTGGCGCTGCGCTCTGAGCGGAGCGGAGCC 4737  
QY 815 CCACAGGAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGCCCAAGTCTCTCGTCC 874  
Db 4738 CGTTGGGCGAGGGTCTCTGGGCGCCACCGCGGCGAGGACGCGTGGACCGAGTCCGCTGTT 4797  
QY 875 CCCCAGGTGCTACT-----GCAGAGAAGATTTGCTCTTCTTAAAGGAAAGGTGTC 925  
Db 4798 CTGTGTGTGTCTACCTGCGCAGACCGCGCGGAGAGGACCACTCTTTGGAGGCTGCTCTC 4857  
QY 926 TGACCTGAGTCTCTC---TGCGTGGTGTGCTGTAAACACAAAGCCAGCTTCCACAT---C 979  
Db 4858 TGGCAGCGCGCTCTCCCACTCCGTGGGCGCGCAGCACACCGCGGCGCGCCCATCCAC 4917  
QY 980 TCTGTGTCAACCCCGCGCAAAATGCTTTCAGCTCAGGCAATTTATTGAGACCAAGCA 1039  
Db 4918 ATCGGCGCACACGCTCCCTGGGACACGCTTGTCCCGGCTGACGCGGAGCAAGCA 4977  
QY 1040 TTTCTTTACTTCCAGGGGAGATGGCCAGAGCGCTCTAAACCCCTCATTTCTTACTCAGCAA 1099  
Db 4978 CTCTCTCTCTCTCAGCGCA---CAAGGAGAGCTGCGGCGCTCTCTTCTTCTCAGCTC 5034  
QY 1100 CTTGAGGCTTACTGCTGGGCGCAGGAGACTGTTGGAGATCATCTTTCTGGGCTCAAG 1159  
Db 5035 TCTGAGGCGGAGCTGCTGGCGCTGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAG 5094  
QY 1160 GCCTAGGACATCAGGACCATCTCTGAGGACACACCGTCTATCCGCTCGATCTGCGCAGAT 1219  
Db 5095 GCTGTGATGCGAGGACTCCCGCAGGTTGCGCGCTGCGCGCTGCGCGCTGCTGCGCAAT 5154  
QY 1220 GCGGCGCTGTTTCAACAGCTGTGTTGAAACCATGAGAGTGCCTATGTCAGATCTCT 1279  
Db 5155 GCGGCGCTGTTTCTGAGCTGTTGGAACACACGCGCAGTGCCTTCTGCGGCTGCTCT 5214  
QY 1280 CAGTCACTTTCAGGTTTCGACACAAACCAAGGTGACAGATGCTCT----- 1329  
Db 5215 CAGACGCACTGCGCGCTGCGAGTGGGTCAACCCAGCAGCGGCTGCTGTGCGCGGGA 5274  
QY 1330 -----TGAAACACAGCCACCGCACCTCAT 1354  
Db 5275 GAAGCCCGAGGCTCTGTGCGCGCGCGGAGGAGGACACAGACCCCGCTGCGCTGT 5334  
QY 1355 GATTTGCTGCGCTGCGAGAGTCTGCGGAGGATATGTTTCTTCTGCGGCTGCTCT 1414  
Db 5335 GAGTGTGCTGCGCAGCAGCAGCGCTGCGGAGGCTGCGGCTGCGGCTGCGGCTGCTCT 5394  
QY 1415 CTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 1474  
Db 5395 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5454  
QY 1475 GAACTTAAAGTTCATCTGTTGGGAAATACGCGCAAGCTATCATCTCAGGAACTGAT 1534  
Db 5455 GAAACCAAGATTCATCTCTGTTGGGAAATACGCGCAAGCTATCATCTCAGGAGCTGAC 5514  
QY 1535 GTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1594  
Db 5515 GTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5574  
QY 1595 CCGGCTGCGAGCAGCGCTGAGGAGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1654  
Db 5575 TCGGCGCGCAGCAGCAGCGCTGCGGTGAGGAGATCTGCGCAAGTCTCTGCTGCTGCTGCT 5634  
QY 1655 GGACACATAGTGTGACAGCTGTTAGGTCATCTTTTACATCAGCAGAGGACATTCAC 1714  
Db 5635 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5694  
QY 1715 GAAGAACAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1774

||||| 5695 AAAGAACAGGCTCTTTTACCGCGCCGAGTGTCTGGAGCAAGTTGCAAGCAATTGGAA 5754  
||||| 1775 CAGGCAACACCTTTGAGAGAGTGCGGCTACGGGAGTGTCTCAGAGAGAGGTGAGGATCA 1834  
||||| 5755 CAGACAGCACTTGAAGAGGGTGCGAGCTGCGGAGCTGTGCGAAGCAGAGGTTCAGGACGA 5814  
||||| 1835 CCAGGACACTGTGTAGCCATGCCATCTGCAGACTGCGCTTCATCCCAAGCCCAAGG 1894  
||||| 5815 TCGGGAAGCCAGGCGCCGCTGTGACGTCCAGACTCCGCTTCATCCCAAGCCCTGACGG 5874  
||||| 1895 CTTGCGGCCCATTTGTGAACATAGTTATAGCATGGGTACAGAGCTTTGGGAGAGGAA 1954  
||||| 5875 GCTGCGGCCGATTGTGAACATGACTACGTCGTGGGAGCCAGAACGTTCCGCGAGAGAAA 5934  
||||| 1955 GCAGGCCAGCATTTACCCAGCGTCTCAGACTCTCTTCAGCATGCTCAACTATGAGCG 2014  
||||| 5935 GAGGGCCGAGCGCTCACCTCGAGGGTGAGGCACTGTTTCAGCGTCTCAACTACGAGCG 5994  
||||| 2015 GACAAAACATCTCACCTTATGGGTCTTCTGTACTGGGTATGAATGATCATCTACAGGAC 2074  
||||| 5995 GCGCGGCGCCCGGCTCTCGGCGCTCTGTCTGGCTGGAGGATATCCACAGGGC 6054  
||||| 2075 CTGCGGCGCCTTTGTGCTGCTGTGCTGTGAGCCAGACACCCAGATGATCTTTGT 2134  
||||| 6055 CTGCGGCACTTCTGCTGCTGTGCGGCGCCAGGACCGCGGCTGTGAGCTGTACTTTGT 6114  
||||| 2135 TAAGGCAATGTGACCGGGGCTTATGATGCATCTCCCGAGGTAAGCTGGTGGAGGTTGT 2194  
||||| 6115 CAAGTGGATGTGACGGGCGGTAGACACCATCTCCCGAGGACAGCTCAGGGAGTAT 6174  
||||| 2195 TGCCAAATGATCAGGCATCTCGAGAGCAGTACTGTATCCGCCAGTATGCAAGTGGTCCG 2254  
||||| 6175 CGCCAGCATCATCAACCCG --- AGAACAGCTACTCGTGGCTGCGTATGCGGTGCCA 6231  
||||| 2255 GAGGATAGCCAGGCCAAGTCCACAGTCTTTAGGAGACAGGTCAACCCCTCTCTGA 2314  
||||| 6232 GAAGCGCCCATCTGGGACGTCGCGAAGGCTTTCAAGAGCCAGCTCTTACCTTTGACAGA 6291  
||||| 2315 CCTCAGCCATACATGGGCGAGTTCTTAAGCATCTGCAGGATTCAGATGCCAGTGCAT 2374  
||||| 6292 CCTCAGCCGTACATGGACAGTTCTGTGGCTCACCTGCAG --- GAGACAGCCCGCT 6345  
||||| 2375 GAGAACTCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAGCAGCCTGTT 2434  
||||| 6346 GAGGGATCGGCTCGTCTATCAGCAGAGCTCTCTCCCTGAATGAGGCGCAGTGGGCTCTT 6405  
||||| 2435 TGACTTCTCTGCACTTCTGCTGCTACAGTGTCTGTAAGATTGGTGAAGGTGCTATAC 2494  
||||| 6406 CGAGCTCTTCTTACGCTTCTATGTGCCACACAGCGCTGCGCATCAGGGCAAGTCTTACGT 6465  
||||| 2495 GCAGTGCAGGGCATCCCGAGGGCTCCAGCCTATCCACCTGCTCTGAGTCTGTGTT 2554  
||||| 6466 CCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCAGCTGCTCTGAGCCTGTGCTA 6525  
||||| 2555 CGGAGACATGGAGAACAGCTGTTCTGAGGTGACGCGGATGGTGTCTTTTACGTTT 2614  
||||| 6526 CGGCGACATGGAGAACAGCTGTTTGGGGATTCGCGGGACGGGCTCTCTCGCTT 6585  
||||| 2615 TGTGTAGCATTTTGTGTGTGAGCGCTCACTTTGGACCAAGCAAAACCTTCTCAGCAC 2674  
||||| 6586 GGTGGATGATTTCTTGTGTGTGACACTCACTACCTCACCCAGCGGAAACCTTCTCAGGAC 6645  
||||| 2675 CTTGGTCCATGGGCTTCTCAGTATGGGTGCATGATAAATTGCGAGAGACAGTGGTGA 2734  
||||| 6646 CTTGGTCCGAGGTGCTCCTCAGTATGGCTGCGTGGTGAACCTTGGGAGACAGTGGTGA 6705  
||||| 2735 CTTCCCTGTGGAGCCTGGTACCTCGTGGTGTGAGTCCATACCACTGCTGCTCACTG 2794  
||||| 6706 CTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGATGCCGCCACCG 6765  
||||| 2795 CTTGTTTCCCTGTTGGCTGTGCTGAGCACTCAGACTTTGGAGGTTGTTCTGTGACTA 2854  
|||||

Db 6766 CTTATTCCTGCTGGTGGGCTGTGCTGGATATCCCGACCCCTGGAGGTGAGAGGACTA 6825  
QY 2855 CTTAGGTTATGCCAGACCTCAATTAAGAAGAGCTCACCTTCAGAGTGTCTTCAAGC 2914  
Db 6826 CTTCAAGCTATGCCCGGACCTCCATCAGAGCCAGTGTCACTTCAACCGCGCTTCAAGGC 6885  
QY 2915 TGGGAGGACCATGCGGAACAGCTCCTCTCGTCTTGGGTGAGGTGTCAGGCTCTATT 2974  
Db 6886 TGGGAGGAAACATGCGTCCGAAACTTTTGGGGTCTTGGGCTGAAGTGTACAGCTGTT 6945  
QY 2975 TCTAGACTTGCAGGTGAACAGCCTCCAGACAGTGTGCATCAATATATACAGATCTTCT 3034  
Db 6946 TCTGGATTGACAGGTGAACAGCCTCCAGACAGTGTGCACCAACATCTACAAGATCTCTCT 7005  
QY 3035 GCTTCAGGCTACAGGTTCCATGATGTGTGATTTCAGTTCCTTTGACCAAGGTGTAG 3094  
Db 7006 GCTTCAGGCGGTACAGGTTTCAGCATGTGTCTGAGCTCCCATTTTCATCAGCAAGTTG 7065  
QY 3095 GAAGAACCTTCATTTCTTCTGGGCATCATCTCCAGCAAGCATCTCTGCTGCTATGCTAT 3154  
Db 7066 GAAGAACCCCATTTTCTTCTGGGCTCATCTCTGACACGGCCTCCCTCTGCTACTCCAT 7125  
QY 3155 CTTGAAGGTCAAGATCCAGGATGACACTAAAG-----GCCTCTGGCTCCTTTCC 3205  
Db 7126 CTTGAAGGCCAAGAACGAGGGATGTCTGCTGGGGGCCAAGGGCGCCGCTCTGCTGCC 7185  
QY 3206 TCTGAAGCGGCACATTTGGCTCTGCTACAGGCTTCTCTGCTCAAGTGGCTGCTCATTC 3265  
Db 7186 CTTCCGAGCGGTGAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCTGACTGACACCG 7245  
QY 3266 TGTCTATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAAATGCTGTGCTCCGAA 3325  
Db 7246 TGTCACTACGTGCCACTCTCTGGGTCACTCAGGACAGCCCAAGCAGCTGAGTCCGAA 7305  
QY 3326 GCTCCAGAGGGGCAATGACCATCTTAAAGCTGACGCTGACCCAGCCCTTAAGCAGAGA 3385  
Db 7306 GCTCCGGGAGCAGCTGACTGCTGCTGGAGGCGGAGCCAAACCCGCACTGCCCTCAGA 7365  
QY 3386 CTTTCAGACCATTTTGGACTAA 3407  
Db 7366 CTTCAAGACCATCTCTGGACTGA 7387

Search completed: February 2, 2004, 07:23:44  
Job time : 1429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

1 on: February 1, 2004, 21:42:51; Search time 6853 Seconds  
(without alignments)  
12398.723 Million cell updates/sec

le: US-09-042-460-1

fect score: 3496

pence: 1 GAATTCGGGTGGAGGCC.....CCGAGTCGTACCAAGCTT 3496

ring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

urched: 22781392 seqs, 12152238056 residues

al number of hits satisfying chosen parameters: 45562784

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

EST:\*

1: em\_estba:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	ID	Description
1	844.8	24.2	851	13 BU702370	BU702370 UI-M-F10-
2	764.8	21.9	851	12 BG917907	BG917907 602820830
3	655.8	18.8	664	13 BQ258274	BQ258274 NISC_kp11
4	612.4	17.5	641	28 AZ972318	AZ972318 2M0246F07

5	513	14.7	599	10	BB618671	BB618671
6	466	13.3	614	10	BB651920	BB651920
7	440.8	12.6	925	12	BM453198	BM453198 AGENCOURT
8	431.2	12.3	866	10	BE371943	BE371943 601217728
9	347	9.9	347	9	AW244516	AW244516 BR_END068
10	302	8.6	326	13	BY149368	BY149368 BY149368
11	272.2	7.8	492	12	BM824748	BM824748 K-EST0096
12	265	7.6	568	4	EX521269	EX521269 R2PD Mus
13	261.8	7.5	534	9	AW318894	AW318894 un09402.y
14	235.6	6.7	389	9	AA281296	AA281296 zt08902.r
15	186.8	5.3	753	13	BU452535	BU452535 603767927
16	172.8	4.9	775	12	BI388013	BI388013 BFL26.002
17	164	4.7	880	13	BU377259	BU377259 603811228
18	160.8	4.6	679	10	BE396606	BE396606 601289077
19	158.8	4.5	610	10	BE514188	BE514188 601316376
20	157.8	4.5	409	9	AA311750	AA311750 EST182469
21	152.2	4.4	715	10	BE396925	BE396925 601290610
22	150	4.3	649	10	BE514070	BE514070 601316575
23	142.6	4.1	696	13	BUI39751	BUI39751 603134527
24	134	3.8	487	28	AZ799615	AZ799615 2M0057107
25	130.2	3.7	668	14	CA380121	CA380121 659344 NC
26	123.2	3.5	779	10	BE268183	BE268183 601125261
27	118	3.4	875	13	BUI22597	BUI22597 603148441
28	109.2	3.1	632	14	CA353864	CA353864 625469 NC
29	104.4	3.0	835	13	BUI11946	BUI11946 603127372
30	101.4	2.9	212	10	BB587267	BB587267 BB587267
31	98	2.8	777	28	AZ826048	AZ826048 2M0101K03
32	97.4	2.8	813	13	BU224024	BU224024 603798349
33	91	2.6	375	9	AA200728	AA200728 mul3h09.r
34	91	2.6	503	9	AI645957	AI645957 mul3h09.y
35	78	2.2	554	28	AQ397020	AQ397020 mgxb0014C
36	76.2	2.2	1230	29	CC211572	CC211572 CH261-14F
37	76.2	2.2	1455	29	CC190951	CC190951 CH261-380
38	76.2	2.2	1554	29	CC190875	CC190875 CH261-38M
39	74.8	2.1	813	10	BGI98331	BGI98331 RST17599
40	66.8	1.9	739	13	EX315053	EX315053 EX315053
41	66.4	1.9	366	10	BF511837	BF511837 UI-H-B14-
42	64.6	1.8	239	12	BI359552	BI359552 OKST5-7(2
43	63.8	1.8	479	12	BM077893	BM077893 pb21ell.y
44	63.2	1.8	519	10	BF802688	BF802688 PM4-CI007
45	63.2	1.8	929	29	CNS06T5X	AL414967 T7 end of

#### ALIGNMENTS

RESULT 1  
BU702370  
LOCUS  
DEFINITION  
UI-M-F10-byx-f-12-0-UI.r1 NIH-BMAP\_F10 Mus musculus cDNA clone  
IMAGE: 6400523 5', mRNA sequence.  
ACCESSION  
BU702370  
VERSION  
BU702370.1  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgc.nci.nih.gov/  
1 (bases 1 to 851)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BNAP)

QY	2818	TCCTGGACACTCAGACTTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCCGACACCTCAA	2877
DB	601	TCCTGGACACTCAGACTTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCCGACACCTCAA	660
QY	2878	TTAAGACGAGCCTCACCTTCCAGAGTGTCTTCAAAGCTGGGAAGACCAATGCGGAACAAGC	2937
DB	661	TTAAGACGAGCCTCACCTTCCAGAGTGTCTTCAAAGCTGGGAAGACCAATGCGGAACAAGC	720
QY	2938	TCCGTGCGGTCTTGGCGTTGAAGTGCAGGTCCTATTCTTAGACTTGCAGCTGCAACAGCC	2997
DB	721	TCCGTGCGGTCTTGGCGTTGAAGTGCAGGTCCTATTCTTAGACTTGCAGCTGCAACAGCC	780
QY	2998	TCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTTCAGGCTACAGGTTCCATG	3057
DB	791	TCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTTCAGGCTACAGGTTCCATG	840
QY	3058	CATGTGTGATT	3068
DB	841	CATGTGTGATT	851
RESULT 2			
BG917907			
LOCUS			
DEFINITION			
602820830F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4949887 5',			
mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLES			
JOURNAL			
COMMENT			
<p>851 bp mRNA linear EST 05-JUN-2001</p> <p>602820830F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4949887 5',</p> <p>mRNA sequence.</p> <p>BG917907</p> <p>BG917907.1 GI:14298383</p> <p>EST.</p> <p>Mus musculus (house mouse)</p> <p>Mus musculus</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</p> <p>Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p> <p>1 (bases 1 to 851)</p> <p>NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a></p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a></p> <p>Tissue Procurement: Jeffrey Green M.D.</p> <p>cDNA Library Preparation: Life Technologies, Inc.</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be</p> <p>found through the I.M.A.G.E. Consortium/LLNL at:</p> <p><a href="http://image.llnl.gov">http://image.llnl.gov</a></p> <p>Plate: LLAM10903 row: k column: 08</p> <p>High quality sequence stop: 753.</p>			
FEATURES			
source			
<p>1..851</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="FVB/N"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:4949887"</p> <p>/sex="female; virgin"</p> <p>/tissue_type="infiltrating ductal carcinoma"</p> <p>/dev_stage="5 months"</p> <p>/lab_host="DH10B"</p> <p>/clone_lib="NCI CGAP Mam6"</p> <p>/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;</p> <p>Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.</p> <p>Library constructed by Life Technologies. Investigator</p> <p>providing samples: Jeffrey Green, M.D., NIH"</p>			
BASE COUNT			
ORIGIN			
<p>184 a 238 c 214 g 215 t</p> <p>21.9%; Score 764.8; DB 12; Length 851;</p> <p>Query Match</p> <p>Best Local Similarity 97.5%; Pred No. 1.2e-196;</p> <p>Matches 819; Conservative 0; Mismatches 17; Indels 4; Gaps 4;</p>			
QY	2582	TGAGGTGCAGCGGGATGGGTTCCTTTTACGCTTTTGTGTGATGACTTCTCTTCGTGACGCC	2641

9 TGGGTGACGCGGATGGG-TGCTTTAGTTGG-TGATGACTTTCTGAGGTGAGCC 66  
 2642 TCACCTGGACCAAGCAAAACCTTCTCAGCACCCTGCTCCATGGCTTCTCAGATAGG 2701  
 67 TCACCTGGACCAAGCAAAACCTTCTCAGCACCCTGCTCCATGGCTTCTCAGATAGG 156  
 2702 GTGCATGATAACTTGCAGACAGAGTGGTGAATCCCTGTTGAGCTGGTACCTGGG 2761  
 127 GTGCATGATAACTTGCAGACAGAGTGGTGAATCCCTGTTGAGCTGGTACCTGGG 186  
 2762 TGCTGAGCTCATACCAAGCTGCTCTCACTGCTGTTTCCCTGGTGGCTTGTCTGCT 2821  
 187 TGCTGAGCTCATACCAAGCTGCTCTCACTGCTGTTTCCCTGGTGGCTTGTCTGCT 246  
 2822 GCACACTGAGCTTTGGAGGTGTTCTGACTACTCAGGTTATGCCAGACCTCAATTA 2881  
 247 GCACACTGAGCTTTGGAGGTGTTCTGACTACTCAGGTTATGCCAGACCTCAATTA 306  
 2882 GACGAGCTTCACTTCCAGAGTGTCTTCAAGCTGGGAGACCATGCGGAACAGCTCT 2941  
 307 GACGAGCTTCACTTCCAGAGTGTCTTCAAGCTGGGAGACCATGCGGAACAGCTCT 366  
 2942 GTCCGTCTTGGCGTTGAAGTGTACGGTCTATTCTAGCTTGCAGTGAACAGCTTCCA 3001  
 367 GTCCGTCTTGGCGTTGAAGTGTACGGTCTATTCTAGCTTGCAGTGAACAGCTTCCA 426  
 3002 GACGCTCTGCATCAATATATACAGATCTTCTGCTTCAGGCTTACAGGTTCCATGCTG 3061  
 427 GACGCTCTGCATCAATATATACAGATCTTCTGCTTCAGGCTTACAGGTTCCATGCTG 486  
 3062 TGTGATTCAGCTTCCCTTTGACCGCGTGTAGGAAGAACTTCACTTTCTGGGCGAT 3121  
 487 TGTGATTCAGCTTCCCTTTGACCGCGTGTAGGAAGAACTTCACTTTCTGGGCGAT 546  
 3122 CATCTCCAGCAGCATCTGCTGATGCTATCTGATGCTTCAAGTCAAGATCCAGGAATGAC 3181  
 547 CATCTCCAGCAGCATCTGCTGATGCTATCTGATGCTTCAAGTCAAGATCCAGGAATGAC 606  
 3182 ACTAAGGCTCTGGCTCTTCTTCCCTGGAAGCGGCACATTTGGCTTGTACAGGCTT 3241  
 607 ACTAAGGCTCTGGCTCTTCTTCCCTGGAAGCGG-ACATTTGGCTTGTACAGGCTT 665  
 3242 CTT-GCTCAGCTGCTGCTCATTTCTGATCTCAAAATGTTCTCTGGGACCTTCAAGGA 3300  
 666 CTTGCTCAGCTGCTGCTCATTTCTGATCTCAAAATGTTCTCTGGGACCTTCAAGGA 725  
 3301 CAGCCCAAAACCTGCTGCGGAGCTCCAGAGGCGACATGACCATCTTAAAGCTG 3360  
 726 CAGCCCAAAACCTGCTGCGGAGCTCCAGAGGCGACATGACCATCTTACAGCTG 785  
 3361 CAGCTGACCCAGCCTTAAGCAGACTTTTACAGCAATTTTGGACTAACCTTGTCTCTTC 3420  
 786 CAGCTGACCCAGCCTTAAGCAGACTTTTACAGCAATTTTGGACTAACCTTGTCTCTTC 845  
 BQ258274 664 bp mRNA linear EST 06-MAY-2002  
 NISC kp11904.q3 Baker mouse embryo e7.5 Mus musculus cDNA clone  
 IMAGE:5409222, mRNA sequence.  
 BQ258274  
 BQ258274.1 GI:20459030  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 664)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 cDNA Library Preparation: J. Baker (Stanford University)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[infoimage.llnl.gov](http://infoimage.llnl.gov)  
 MGI:1845958

Plate: L1AM12043 row: N column: 7

Seq primer: Sp6 primer.

## FEATURES

## source

1..664  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:5409222"  
 /tissue\_type="embryo, late gastrula"  
 /dev\_stage="embryo, 7.5 dpc"  
 /lab\_hosts="XLI-Blue"  
 /clone\_lib="Baker mouse embryo e7.5"  
 /note="vector: pCS105; Site 1: NotI; Site 2: SalI; cDNA  
 made by oligo-dT priming. Directionally cloned into  
 SalI/NotI sites using the following 5' adaptor:  
 5'-TCGACCCAGCCTCCG-3', size-selected for average insert  
 size 1.8-1.9 kb. Library constructed by J. Baker (Stanford  
 University)."

BASE COUNT 155 a 177 c 183 g 149 t

## ORIGIN

Query Match 18.8%; Score 655.8; DB 13; Length 664;  
 Best Local Similarity 99.7%; Pred. No. 4.8e-167;  
 Matches 557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1103 CCAGCTTAACCTTGAATGGGCGCCAGGAGACTGTGGAGATCATCTTTCTGGGTCAAGGCC 1162  
 Db 6 CGAGCTTAACCTTGAATGGGCGCCAGGAGACTGTGGAGATCATCTTTCTGGGTCAAGGCC 65  
 QY 1163 TAGACATCAGGACCACTCTGCAGGACACACCGTCTATCGCTCGATCTGCAGATGCG 1222  
 Db 66 TAGACATCAGGACCACTCTGCAGGACACACCGTCTATCGCTCGATCTGCAGATGCG 125  
 QY 1223 GCCCTGTTCACACAGCTGCTGGTGAACCATGACAGAGTGCCATATGTCTCAGACTCTCTCAG 1282  
 Db 126 GCCCTGTTCACACAGCTGCTGGTGAACCATGACAGAGTGCCATATGTCTCAGACTCTCTCAG 185  
 QY 1283 GTCACTATGACGTTTCGAACAGACAAACAGAGTGACAGATGCTTGAACACCGGCC 1342  
 Db 186 GTCACTATGACGTTTCGAACAGACAAACAGAGTGACAGATGCTTGAACACCGGCC 245  
 QY 1343 ACCGCACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGCGAGGTATATGTTTCT 1402  
 Db 246 ACCGCACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGCGAGGTATATGTTTCT 305  
 QY 1403 TCGGCTGTCTCTGCAAGTGTGTCTGCTAGTCTCTGGGTACACAGGCAATGAGCG 1462  
 Db 306 TCGGCTGTCTCTGCAAGTGTGTCTGCTAGTCTCTGGGTACACAGGCAATGAGCG 365  
 QY 1463 CCGTCTCTTTAAGAACTTAAGAAAGTTTCATCTGTTGGGAAATACGCAAGCTATCACT 1522  
 Db 366 CCGTCTCTTTAAGAACTTAAGAAAGTTTCATCTGTTGGGAAATACGCAAGCTATCACT 425  
 QY 1523 GCAGGAACCTGATGGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGCAGGCCCGG 1582  
 Db 426 GCAGGAACCTGATGGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGCAGGCCCGG 485  
 QY 1583 GAAGGACCGTGTCCCGCTGCAGACACCGTCTGAGGAGAGAGTCTGGCTACCTTCT 1642  
 Db 486 GAAGGACCGTGTCCCGCTGCAGACACCGTCTGAGGAGAGAGTCTGGCTACCTTCT 545  
 QY 1643 GTTCTGCTGATGGACACATACGTGTACAGCTGCTTAGGTCACTTTTACATCACAGA 1702

## SULT 3

## 258274

## FINITION

## SESSION

## SION

## WORDS

## RCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL



|||||  
545 GTTCTGGCTGATGGACACATACGTGGTACAGCTGTTAGGTCACTTTTACATCACAGA 605  
|||||  
1703 GAGCACATCCAGAGACAGGCTCTTCTTACCGTAAGAGTGTGGAGCAGCTGC 1761  
|||||  
606 GAGCACATCCAGAGACAGGCTCTTCTTACCGTAAGAGTGTGGAGCAGCTGC 664  
|||||

SULT 4  
972318  
TUS  
PINITION  
AZ972318 641 bp DNA linear GSS 27-APR-2001  
2M0246F07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
clone UUGC2M0246F07 F, genomic survey sequence.  
SESSION  
AZ972318  
RSION  
AZ972318.1 GI:13843545  
YWORDS  
GSS.  
URCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 641)  
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

Unpublished  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0246 row: F column: 07  
Seq primer: CGTTGTAAACGACGACCCAGT  
Class: plasmid ends  
High quality sequence stop: 641.  
Location/Qualifiers  
1. 641  
/organism="Mus musculus"  
/mol type="genomic DNA"  
/strain="C57BL/6J"  
/db xref="taxon:10090"  
/clone="UUGC2M0246F07"  
/sex="female"  
/lab host="E. coli strain XL10-Gold, Tl-resistant, F-"  
/clone lib="Mouse 10kb plasmid UUGC2M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (female) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (GI|4732114|gb|AP19072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adaptor vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ATUES  
source

SE COUNT  
IGIN

Query Match 17.5% Score 612.4; DB 28; Length 641;  
Best Local Similarity 99.0%; Pred. No. 3.1e-155;  
Matches 616; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 762 CTTTCAGCTAAGAGGCGCAGATGCTATCTCTGTCGAGAGTGGAGGAGGCCCCACAGG 821  
DB 20 CATGAATCGCGAAGGCGCAGATGCTATCTCTGTCGAGAGTGGAGGAGGCCCCACAGG 79  
QY 822 CAGTGTGTACCAACCCCATCAGGCAATCATGGTCCCAAGTCCCTGCTCGTCCCGGAG 881  
DB 80 CAGTGTGTACCAACCCCATCAGGCAATCATGGTCCCAAGTCCCTGCTCGTCCCGGAG 139  
QY 882 GTGCTACTGTCAGAGGAGAGATTTGTTCTTCTTAAGAGGAGGTGCTGACCTGAGTCTCT 941  
DB 140 GTGCTACTGTCAGAGGAGAGATTTGTTCTTCTTAAGAGGAGGTGCTGACCTGAGTCTCT 199  
QY 942 GGGTCGGTGTGCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACACCCCGCCAA 1001  
DB 200 GGGTCGGTGTGCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACACCCCGCCAA 259  
QY 1002 AATGCTTTTTCAGCTCAGGCCATTTATTGAGACCAAGCATTTCTTTTACTCCAGGGAGAT 1061  
DB 260 AATGCTTTTTCAGCTCAGGCCATTTATTGAGACCAAGCATTTCTTTTACTCCAGGGAGAT 319  
QY 1062 GGCACAAGAGCGCTTAAACCCCTCATTTCTACTCAGCAACCTCCAGCCTTAACCTTGCTGG 1121  
DB 320 GGCACAAGAGCGCTTAAACCCCTCATTTCTACTCAGCAACCTCCAGCCTTAACCTTGCTGG 379  
QY 1122 GCCAGGAGACTGTGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCACTC 1181  
DB 380 GCCAGGAGACTGTGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCACTC 439  
QY 1182 TGCAGGACACACCGCTCTATCGCTGCATATCTGGCAGATGGCGCCCTGTTTCCAAACAGCTG 1241  
DB 440 TGCAGGACACACCGCTCTATCGCTGCATATCTGGCAGATGGCGCCCTGTTTCCAAACAGCTG 499  
QY 1242 CTGTGTGAACCATCAGAGTGCCTCAATATGTGAGACTCTCTAGGTCAATTCAGGATTCGA 1301  
DB 500 CTGTGTGAACCATCAGAGTGCCTCAATATGTGAGACTCTCTAGGTCAATTCAGGATTCGA 559  
QY 1302 ACAGCAAAACCAAGGTGACAGATGCTTGAACACCAAGCCACCGACCTCATGGATTTG 1361  
DB 560 ACAGCAAAACCAAGGTGACAGATGCTTGAACACCAAGCCACCGACCTCATGGATTTG 619  
QY 1362 CTCCTCGCTGTCAGCAGCTCCCT 1383  
DB 620 CTCCTCGCTGTCAGCAGCTCCCT 641

RESULT 5  
BB618671  
LOCUS  
BB618671 RIKEN full-length enriched, 8 days embryo Mus musculus  
cDNA clone 5730412M20 5', mRNA sequence.

DEFINITION  
BB618671  
VERSION  
BB618671.1 GI:16458173  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)

REFERENCE  
AUTHORS  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
Hiranoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda  
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,J., Ohno,M., Sakai  
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki  
D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
Unpublished  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwake, S., Inoue, K., Togawa, M., Ohaza, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
a mouse tissue

```

SE COUNT      91 a 185 c 181 g 142 t
IGIN
Query Match    14.7%; Score 513; DB 10; Length 599;
Best Local Similarity 100.0%; Pred.No. 3.5e-128;
Matches 513; Conservative 0; Mismatches -0; Indels 0; Gaps 0;

10 GTGGAGGCCCATCCGGCCTTGAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGC 69
86 GTGGAGGCCCATCCGGCCTTGAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGC 145
70 GCTCTCTGTGCGCACCGCATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGGCC 129

```

146	GCTCTCTGCTGCGCAGCCGATACCGGGAGGGTGTGGCCGTGGCAACCTTTGTGCGCGCC	205
130	TGGGCCCCGAGGGCAGGCGGCTTGTGCAACCCGGGGAGCCCGAAGATCTTACCGCACTTTGG	189
206	TGGGCCCGAGGGCAGGCGGCTTGTGCAACCCGGGGAGCCCGAAGATCTTACCGCACTTTGG	265
190	TTGCCCAATGCTAGTGTGATGCATCTGGGGCTCAAGGCTTCCACCTGCCAGCCTTTCCT	249
266	TTGCCCAATGCTAGTGTGATGCATCTGGGGCTCAAGGCTTCCACCTGCCAGCCTTTCCT	325
250	TCCACCAAGTGTCACTCCTCGAAGAGCTCGTGGCCAGGGTGTGCGAGAGACTCTGCCAGC	309
326	TCCACCAAGTGTCACTCCTCGAAGAGCTCGTGGCCAGGGTGTGCGAGAGACTCTGCCAGC	385
310	GCAACGAGAGAAACGTGCTGGCTTTTGGCTTTTGAAGTCTCTTAAACGAGGCCAGAGCGCGGC	369
386	GCAACGAGAGAAACGTGCTGGCTTTTGGCTTTTGAAGTCTCTTAAACGAGGCCAGAGCGCGGC	445
370	CTCCCATGGCTTTCACCTAGTAGGTCGGTAGCTATTGCCCAACACATGTTATTGAGACCC	429
446	CTCCCATGGCTTTCACCTAGTAGGTCGGTAGCTATTGCCCAACACATGTTATTGAGACCC	505
430	TGCGTGTCAAGTGTGCATGATGACTACTCTTGAGCCGAGTGGCGGACGACCTGCTGTCT	489
506	TGCGTGTCAAGTGTGCATGATGACTACTCTTGAGCCGAGTGGCGGACGACCTGCTGTCT	565
490	ACCTGCTGGCACATGTGCTCTTTATCTTCTGG	522
566	ACCTGCTGGCACATGTGCTCTTTATCTTCTGG	598

RESULT 6

BB851920

LOCUS

DEFINITION

BB851920

BB851920 RIKEN full-length enriched, ES cells Mus musculus cDNA

clone C330020G14 5', mRNA sequence.

614 bp

mRNA

linear

EST 26-OCT-2001

RESULT 6	BB651920	614 bp	mRNA	linear	EST 26-OCT-2001
BB651920	BB651920	RIKEN full-length enriched, ES cells	Mus musculus	CDNA	
LOCUS	clone C330020G14 5', mRNA sequence.				
DEFINITION					
ACCESSION	BB651920.1	GI:16486058			
VERSION	BB651920				
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 614)				
AUTHORS	Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiranoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, Y., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.				
TITLE	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)				
JOURNAL	Unpublished				
COMMENT	Contract: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> 10 (10), 1617-1630 (2000) wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i>				





Email: rancourt@ucalgary.ca; URL: http://www.acs.ucalgary.ca/rancourt

DNA sequencing by: University Core DNA Services, University of Calgary. Submitted sequence has been trimmed at both ends to remove the adaptor oligos containing the EcoRI sites; i.e. GAATTCGGACTA (beginning) and TAGTCGAATTC (end) removed. Therefore, reported insert length is longer than actual EST sequence length.  
Insert length: 359 Std Error: 10.00  
Seq primer: T3 Or T7.

# FEATURES

Location/Qualifiers  
1. 347  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129"  
/db\_xref="taxon:10090"  
/clone="06B09"  
/cell\_type="embryonic stem (ES) cell"  
/cell\_line="D3"

/lab\_host="DH5 alpha"  
/clone\_lib="Bain Rancourt retinoic acid induced ES cell neural differentiation subtraction library"  
/note="Vector: pBluescript II SK+ (Stratagene); Site 1: EcoRI; Site 2: EcoRI; Library constructed by Dr. Gerard Bain (present address: Hoechst-ARIAD Genomic Center, ARIAD Pharmaceuticals Inc., 26 Lansdowne Street, Cambridge, Massachusetts, 02139-4234, U.S.A.). To isolate cDNAs corresponding to mRNAs which are upregulated during the neural differentiation of ES cells in vitro, the subtractive hybridization technique of Wang and Brown [1] was employed. Poly(A)+ RNA was prepared from both undifferentiated ES cells and from embryoid bodies which had been cultured for 4 days in the absence of RA followed by an additional 3 days in the presence of 0.5 (M RA (4-/3+ cells). These poly(A)+ RNAs were converted to double-stranded cDNA using the Superscript Choice System (Gibco). Aliquots of both cDNAs were digested with the restriction enzymes AluI and AluI plus RsaI. An adaptor oligo [1] containing an EcoRI site was ligated to the ends of the restricted cDNAs to provide primer binding sites and large amounts of each cDNA population were then produced by the polymerase chain reaction (PCR) as described [1]. Amplified cDNA from undifferentiated ES cells was biotinylated using Photoprobe biotin (Vector Laboratories) according to the manufacturer's protocol. 2.5 ug of amplified cDNA from 4-/3+ cells was mixed with 50 ug of biotinylated ES cell cDNA, denatured by boiling, and hybridized for 20 h. Double stranded cDNAs containing biotin were removed by streptavidin/phenol treatment as described [1]. The remaining subtracted cDNA was mixed with an additional 25 mg of biotinylated ES cell cDNA, denatured by boiling, and hybridized for 2 h. The remaining cDNA was amplified by PCR [Wang and Brown, 1991] exactly as described above. The cDNA obtained from this subtraction procedure was digested with EcoRI and ligated into pBS II SK+ (Stratagene) followed by transformation into E. coli DH5 cells. Individual colonies were picked and the corresponding plasmids were isolated either by an alkaline lysis miniprep procedure [2], or using the Qiaprep spin miniprep kit (Qiagen). Sequence analysis was performed using the Big Dye Cycle Sequencing kit and an ABI373 sequencer (University Core DNA Services, University of Calgary). 1. Wang Z., Brown DD (1991) A gene expression screen. Proc. Natl. Acad. Sci. USA 88, 11505-11509. 2. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York."

78 a 88 c 100 g 81 t

ASE COUNT  
RIGIN

Query Match 9.9%; Score 347; DB 9; Length 347;  
Best Local Similarity 100.0%; Pred. No. 4.2e-83;

Matches	347;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	2227	ACTGATATCGCCAGTATGCACTGCTCCGGAGAGATAGCCAAAGCCAAAGTCCT	2286						
Db	347	ACTGATATCGCCAGTATGCACTGCTCCGGAGAGATAGCCAAAGCCAAAGTCCT	288						
QY	2287	TTAGGAGACAGTACACCACTCTCTGACCTCCAGCCATACATGGGCCAGTTCTTAAAG	2346						
Db	287	TTAGGAGACAGTACACCACTCTCTGACCTCCAGCCATACATGGGCCAGTTCTTAAAG	228						
QY	2347	ATTCGAGGATTCAGATCCAGTGCACCTGAGGAACCTCGTTGTGTCATCGAGCAGCATCT	2406						
Db	227	ATTCGAGGATTCAGATCCAGTGCACCTGAGGAACCTCGTTGTGTCATCGAGCAGCATCT	168						
QY	2407	CTATGATGAGCAGCAGCAGCCTGTTGACTTCTTCCCTCCACTTCTCGGTACACAGTG	2466						
Db	167	CTATGATGAGCAGCAGCAGCCTGTTGACTTCTTCCCTCCACTTCTCGGTACACAGTG	108						
QY	2467	TGCTAAAGATTGGTCACAGGTGCTATACGCACTGAGGATCCAGGGCATCCCCAGGGCTCCAGCC	2526						
Db	107	TGCTAAAGATTGGTCACAGGTGCTATACGCACTGAGGATCCAGGGCATCCCCAGGGCTCCAGCC	48						
QY	2527	TATCCACCTCTGCTGCACTGCTGTTTCCGAGACATGGAGAACAG	2573						
Db	47	TATCCACCTCTGCTGCACTGCTGTTTCCGAGACATGGAGAACAG	1						

## RESULT 10

BY149368 326 bp mRNA linear EST 10-DEC-2002  
LOCUS  
DEFINITION  
Mus musculus cDNA clone L930261L24 5', mRNA sequence.  
ACCESSION  
BY149368  
VERSION  
BY149368.1 GI:26285897  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
1 (bases 1 to 326)  
AUTHORS

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yegi, K., Tonaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojbori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, B.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltas, L., Marchionni, L., McKenzie, L., Mikl, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takanaka, Y., Taylor, X.S., Teasdale, R.D., Tomita, M., Varardo, R., Wagner, L., Wanlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, T., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)  
JOURNAL  
MEDLINE  
PUBMED  
1246851  
COMMENT  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp/  
URL: http://genome-gsc.riken.go.jp/  
Alzawa, K., Akimura, T., Arakawa, T.,  
'T', Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,  
'Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,  
'Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
'M., Waki, K., Watanabe, A., Yamamoto, M., and Hayashizaki, Y. Direct  
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN Integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site (http://genome-gsc.riken.go.jp) for  
further details.

## FEATURES

## Source

Location/Qualifiers  
1..326  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="J930261124"  
/class="whole body"  
/dev\_stage="17.5 days embryo"  
/clone\_1db="RIKEN full-length enriched, 17.5 days embryo  
whole body"

## BASE COUNT

58 a 92 c 98 g 78 t

Query Match 8.6%; Score 302; DB 13; Length 326;  
Best Local Similarity 98.1%; Pred. No. 7.2e-71; Indels 1; Gaps 1;  
Matches 316; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

236 TGCCGACCTTCTCCACAGGAGTGCATCCCTGAAAGAGCTGTGGCCAGGTTGTGCA 295  
1 TGCCGACCTTCTCCACAGGAGTGCATCCCTGAAAGAGCTGTGGCCAGGTTGTGCA 60  
296 GAGACTCTGCGAGCGCAACGAGAAAGCTGTGGCTTTTGGCTTGAAGCTGTTAACA 355  
61 GAGACTCTGCGAGCGCAACGAGAAAGCTGTGGCTTTTGGCTTGAAGCTGTTAACA 120  
356 GGGCAGAGGGGGGCTCCATGCTCCTGCTAGTGAAGTGGTGAAGTGAAGTGAAGTGA 414  
121 GGGCAGAGGGGGGCTCCATGCTCCTGCTAGTGAAGTGGTGAAGTGAAGTGAAGTGA 180  
415 CTGTTATTTAGACCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 474  
181 CTGTTATTTAGACCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 240  
475 AAGACCTGCTGTCTACTCTGTGGCAACGCTGTTTATCTTTGAGGCCGCCCAAGT 534  
241 AAGACCTGCTGTCTACTCTGTGGCAACGCTGTTTATCTTTGAGGCCGCCCAAGT 300  
535 GTGCTTACAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556  
301 GTGCTTACAGAGGAGATGGCC 322

RESULT 11  
BM824748 492 bp mRNA linear EST 06-MAR-2002  
LOCUS K-EST0096335 S22SNUI6n1 Homo sapiens cDNA clone S22SNUI6n1-99-E07  
DEFINITION 5', mRNA sequence.

ACCESSION BM824748  
VERSION BM824748  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 492)  
Oh, K.Y., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr

High quality sequence stop: 492.  
Plate: 99 row: E column: 07  
Location/Qualifiers  
1..492

## FEATURES

## Source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNUI6n1-99-E07"  
/sex="F"  
/tissue\_type="Asclites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/lab\_1db="S22SNUI6n1"  
/note="Organ: Stomach; Vector: pT73-Pac; Site: 1; EcoRI;  
Site 2: NotI; The S22SNUI6 library was contributed by the  
Seared laboratory and it was constructed as described by  
Bonald, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
Research 6(9): 791-806. RNA was prepared from harvested  
cells of SNU-16 culture. SNU-16 cell was obtained from  
Korean Cell Line Bank (KCLB). SNU-16 was established from  
ascitic fluids of Korean patients by Park J.G. et al.  
(1990), Cancer Res 50: 2773-2780."

## BASE COUNT

96 a 152 c 131 g 113 t

Query Match 7.8%; Score 272.2; DB 12; Length 492;  
Best Local Similarity 75.7%; Pred. No. 1.2e-62;  
Matches 337; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

2730 GTGAATCTCCCTGTGAGAGCTGTGATCCCTGGGTGTGAGCTCATACAGCTGCTCT 2789  
1 GTGAATCTCCCTGTGAG 60  
2790 CACTGCTGTTTCCCTGT 2849  
61 CAGGGCTATTTCCCTGT 120  
2850 GACTACTCAGGTATGAGCCAGACCTCAATTAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
121 GACTACTCAGGTATGAGCCAGACCTCAATTAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
2910 AAGCTGGAG 2969  
181 AAGCTGGAG 240  
2970 CTATTCTAGACTTGGAG 3029







REFERENCE 1 (bases 1 to 753)  
 Boardman, P.E., Samz-Ezquerro, J., Overton, I.M., Burt, D.W., Boesch, E.,  
 Fong, W.T., Tickie, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomedical Sciences  
 University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QP, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES  
 source location/Qualifiers

1..753  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer"  
 /db\_xref="taxon:9031"  
 /clone="CHRST696f7"  
 /sex="Female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEORNI4"  
 /note="Organ: ovary; Vector: Bluescript II KS(+); Site\_1:  
 EcoRI; Site\_2: NotI; This normalized library was  
 constructed from 1 million independent clones. cDNA  
 synthesis was initiated using an oligo(dT) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunt-ended, ligated to NotI adapters, digested with EcoRI  
 , size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."

BASE COUNT 230 a 155 c 132 g 236 t  
 ORIGIN

Query Match 5.3%; Score 186.8; DB 13; Length 753;  
 Best Local Similarity 62.7%; Pred. No. 2.8e-39;  
 Matches 312; Conservative 0; Mismatches 177; Indels 9; Gaps 1;

Db 2919 AAGACCATGCGAACAAGCTCCTGCGGTGCGGTGAAGTGCAGGTCATTTCTA 2978  
 5 AAAAAATGAAATGCAATTAAGTCAAGTCCCAAGTGAATGCCATCCTTACTT 64  
 QY 2979 GACTTCAGGTGAACAGCTCCAGACAGTCTGCATCATATATACAAATCTTCTGCTT 3038  
 65 GACTTAAAGATCAACAGCTTCAGACAGTCTTAATTAACATCTACAAATATTTTACTT 124  
 Db 3039 CAGGCTACAGGTTCCATGATGATTCAGTTCCTTCCCTTGAACAGCTGTAGGAAG 3098  
 125 CAGGCTTACAGTTCATGCTGTGTCTTCACTTCATTCACACAAAGTTAGGAAT 184  
 QY 3099 AACCTCACATCTTTTGGGACATCTCCAGCAAGCATCTGCTGTATGCTATCTTG 3158  
 185 AATCTGATTTCTTCTTAAGATCATCTGTGATGCTTCATGCTCATTTTATCTTG 244  
 Db 3159 AAGTCAAGATCCAGAAATGACACT-----AAAGGCTCTGGCTCCTTCTCTCT 3209  
 245 AAAGCTAAATCAAGAGTTCTTTAGTAGCAAAAGATGATCTGTATGTTCCCTTTT 304  
 QY 3210 GAAGCCGACATGCTGTGCTACAGAGCTTCTGCTCAAGCTGCTGCTCATTTCTGTC 3269  
 305 GAGGACAGCAAAATGGCTGTGCTACATGCTTCATGTCAAACTGTCCAAACCAAGTT 364  
 QY 3270 ATGTACAAATGTCTCTGGAGACCTGTGAAGACAGCCCAAAACTGTGCTGCGGAAGCTC 3329

Db 365 ATTACAAATGCTTACTTAGCCCTTAAGTCTAATGATGCATCTGTTGGAGATC 424  
 QY 3330 CCAGAGCGACATGACCATCTTAAAGCTGAGCTGACCCCTTAGACACACACTTT 3389  
 Db 425 CCAAGGATCTATGAACTGCTGAAGAGCGGTGACGAAACCATCGCTTTGTCAAGATTTC 484  
 QY 3390 CAGACATTTTGACTTA 3407  
 Db 485 AAACATATGAGACTTA 502

Search completed: February 2, 2004, 03:44:55  
 Job time : 6855 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 1, 2004, 15:05:59 ; Search time 9856 Seconds  
(without alignments)  
4657.123 Million cell updates/sec

Title: US-09-042-460-2  
Sequence: 1 MTRAPRCPRVRSLSRSRYRE.....TLKAAADPALSTDFQTILD 1122

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n model -DEV=xlh  
-Qc/cgn2.1/usfpo.spool/us09042460/runat.01022004.150551.17014/app.query.fasta\_1.1287  
-DB=genemb1 -QFMT=fastap -SURFIX=rgs -WINDMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=45 -MODE=LOCAL  
-OUTFMT=psc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09042460 @CGN 1.1 6581 @runat.01022004.150551.17014 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmb1:.\*  
1: gb\_ba:.\*  
2: gb\_hcg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pac:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_str:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vi:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pac:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_str:.\*  
28: em\_un:.\*

29: em\_vi:.\*  
30: em\_hcg\_hum:.\*  
31: em\_hcg\_inv:.\*  
32: em\_hcg\_other:.\*  
33: em\_hcg\_mus:.\*  
34: em\_hcg\_pin:.\*  
35: em\_hcg\_rtd:.\*  
36: em\_hcg\_mam:.\*  
37: em\_hcg\_vrt:.\*  
38: em\_sy:.\*  
39: em\_higo\_hum:.\*  
40: em\_higo\_mus:.\*  
41: em\_higo\_other:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5901	100.0	3369	10 AF073311	AF073311 Mus muscu
2	5901	100.0	3426	10 AF051911	AF051911 Mus muscu
3	4751	80.5	4170	10 AF149012	AF149012 Mesocric
4	3505	59.4	3396	6 BD081553	BD081553 Adult bon
5	3505	59.4	3396	6 BD094749	BD094749 The cell
6	3505	59.4	3396	6 BD096291	BD096291 Cells cap
7	3505	59.4	3399	6 AX481414	AX481414 Sequence
8	3505	59.4	3399	6 AR104587	AR104587 Sequence
9	3505	59.4	4015	6 AR175848	AR175848 Sequence
10	3505	59.4	4015	6 AR182221	AR182221 Sequence
11	3505	59.4	4015	6 AR224455	AR224455 Sequence
12	3505	59.4	4015	6 AR226390	AR226390 Sequence
13	3505	59.4	4015	6 AR243328	AR243328 Sequence
14	3505	59.4	4015	6 AR263555	AR263555 Sequence
15	3505	59.4	4015	6 AR265996	AR265996 Sequence
16	3505	59.4	4015	6 AX019310	AX019310 Sequence
17	3505	59.4	4015	6 AX131979	AX131979 Sequence
18	3505	59.4	4015	6 AX552695	AX552695 Sequence
19	3505	59.4	4015	6 BD011044	BD011044 Human tel
20	3505	59.4	4015	6 BD011832	BD011832 Quantitat
21	3505	59.4	4015	6 BD082985	BD082985 Method fo
22	3505	59.4	4015	6 BD131727	BD131727 Method fo
23	3505	59.4	4015	6 E36793	E36793 Human telom
24	3505	59.4	4015	9 AF015950	AF015950 Homo sapi
25	3505	59.4	4042	6 AX001446	AX001446 Sequence
26	3505	59.4	4042	6 AX003121	AX003121 Sequence
27	3505	59.4	4042	6 BD136185	BD136185 Human tel
28	3505	59.4	4070	6 AX391846	AX391846 Sequence
29	3498	59.3	4027	9 AF018167	AF018167 Homo sapi
30	3498	59.3	8960	12 AF043739	AF043739 Synthetic
31	3496	59.2	4037	6 BD011070	BD011070 Human tel
32	3496	59.2	4037	6 E36819	E36819 Human telom
33	3496	59.2	13766	6 AX553919	AX553919 Sequence
34	3473	58.9	8742	6 AX553924	AX553924 Sequence
35	3287	55.7	4029	6 AR104586	AR104586 Sequence
36	3287	55.7	4029	6 AR175847	AR175847 Sequence
37	3287	55.7	4029	6 BD011069	BD011069 Human tel
38	3287	55.7	4029	6 E36818	E36818 Human telom
39	3276.5	55.5	3855	6 AR243330	AR243330 Sequence
40	3276.5	55.5	3855	6 BD011046	BD011046 Human tel
41	3276.5	55.5	3855	6 E36795	E36795 Human telom
42	3217.5	50.9	3210	6 AB085628	AB085628 Homo sapi
43	3002	50.9	1758	10 AF247818	AF247818 Rattus no
44	2989	50.7	3028	9 AB086379	AB086379 Homo sapi
45	2919	49.5	2992	9 AB086950	AB086950 Homo sapi

RESULT 1

#### ALIGNMENTS

AF073311 3369 bp mRNA linear ROD 09-SRP-1998  
 LOCUS AF073311  
 DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.  
 ACCESSION AF073311  
 VERSION AF073311.1 GI:3551846  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 3369)  
 AUTHORS Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.  
 TITLE Expression of mouse telomerase catalytic subunit in embryos and adult tissues  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)  
 MEDLINE 98393668  
 PUBMED 9724727  
 REFERENCES 2 (bases 1 to 3369)  
 AUTHORS Martin-Rivera, L., Herrera, E. and Blasco, M. A.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of Biotechnology, Cantablanco, Madrid 28049, Spain  
 FEATURES  
 source location/qualifiers  
 1..3369  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1..3369  
 /codon\_start=1  
 /product="telomerase catalytic subunit"  
 /protein\_id="AAC34821.1"  
 /db\_xref="GI:3551847"

## CDS

BASE COUNT 737 a 959 c 891 g 782 t  
 ORIGIN  
 AOKLRCRLPAAITITLKAADPLSLDFTOTILD"

## Alignment Scores:

Pred. No.: 1.63e-299 Length: 3369  
 Score: 5901.00 Matches: 1122  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-042-460-2 (1-1122) x AF073311 (1-3369)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20  
 Db 1 ATGACCCCGGCTCCCTTGGCCCGGCTGGCTCTCTGCGCCAGCCATCCCGGAG 60  
 QY 21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgGluValGln 40  
 Db 61 GTGTGGCCGCTGGCACTTTGTGGGGGCTTGGGGCCCGAGGCGAGCGGCTTGTGCA 120

QY 41 ProGlyAspProIleIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisIleP 60  
 Db 121 CCGGGGAGACCCGAGAGATCTACCGCACTTGGTGGCCAAATGCTAGTGTGATGACG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHISGlnValSerSerLeuLysGluLeu 80  
 Db 181 GGCTCAAGCCTCCACCTCCGACCTTCTTCCACAGGTGTCTATCCCTGAAAGAGCTG 240  
 QY 81 ValAlaArgValValGlnArgLeuGlyGluValArgGlnValArgAspValLeuAlaPheGly 100  
 Db 241 GTGGCAGGAGTGTGAGAGACTTGGCCAGCCGACAGAGAAAGTGTGCTTGTGGC 300  
 QY 101 PheGluLeuLeuGlnValAlaArgGlyGlyProPheValAlaPheThrSerSerValArg 120  
 Db 301 TTGAGGCTGTTAAGAGGACAGAGGCGGGGCTCCCATGGCTTCACTAGTACGCTGGT 360  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaThrMetLeuLeu 140  
 Db 361 AGCTACTTGGCCAACTGTATTGAGACCTTGGCTGTGTCAGTGTGCAATGAGTACTG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 421 TTGAGCCGAGTGGGAGAGACAGCTGCTGCTACCTGCTGGACACTGTCCTTATCTT 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysVal 180  
 Db 481 CTGTGTGCCCCAGCTGTGTGCTACCAAGGTGTGGGTCTCCCTGTACCAATTTGTGCC 540  
 QY 181 ThrThrAspIleThrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 541 ACCAGGATTTCTGGCCCTCTGTGTCCGTAGTACAGGCCACCCGACCCGTGGCAGG 600  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnValPro 220  
 Db 601 AATTCACTAATCTTGTAGTCTTCAACAAGACCAAGAGAGTGTGCGACAGAGACACCG 660  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 661 AAACCTTGGCTTGCATCTGATCTCAAGGTACAAAGAGGACTGAGTCTCCAGTCAAGT 720  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyProHis 260  
 Db 721 GTGCTTCACTAAGAGAGGCAATGCTATCTGTGCCAGAGTGTGAGGAGGAGCCAC 780  
 QY 261 ArgGlnValLeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerPro 280  
 Db 781 AGGAGGTGTACCAACCCATGAGGCAATATCATGGGTGCCAAGTCTGTGCTGCCCC 840  
 QY 281 GluValProThrAlaGluLysAspLeuSerSerLysGlyLysValSerAspLeuSerLeu 300  
 Db 841 GAGGTGCTACTGACAGAAAGATTTGTCTTCAAGAAAGGTGTGAGCTGAGTCTC 900  
 QY 301 SerGlySerValCysCysLysHisLysProSerSerThrSerLeuLeuSerProArg 320  
 Db 901 TCTGGTGTGGTGTGTAAACCAAGCCAGCTCACAATCTCTGTGCAACACCCGCG 960  
 QY 321 GluAsnAlaPheGlnLeuArgProPheIleGluThrArgHisPheLeuTyrSerArgGly 340  
 Db 961 CAAATCTCTTCAAGCTCAGGCACTTATTTAGACCAAGCACTTCTTATCCAGGGGA 1020  
 QY 341 AspGlyGlnValArgLeuLeuProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThr 360  
 Db 1021 GATGGCCAAAGCGTCTAAACCCCTCATCTCACTCAGCAACTCCAGCTCAAGCTGACT 1080  
 QY 361 GlyAlaArgArgLeuValGluIleIlePheLeuGlySerArgProArgThrSerGlyPro 380  
 Db 1081 GGGGCCAGAGACTGTGGAGATCACTTTCTGGGTCAAGGCTTAAAGCATCAGAGCA 1140  
 QY 381 LeuCysArgThrHisArgLeuSerArgArgTyrTyrGlnMetArgProLeuPheGlnGln 400  
 Db 1141 CTCTGAGAGACACACGCTGTATGGCTCATATCTGCAATGCGAGCGCCCTGTTCACAG 1200  
 QY 401 LeuLeuValAsnHisAlaGluCysGlnTyrValArgLeuLeuArgSerHisCysArgPhe 420

```
Db 1201 CTGCTGTTGAACATGAGAGTGCACATATGTCAGACTCTCTCAGGTGCATTCGACGTTT 1260
Qy ArgThrAlaAsnGlnGlnValThrAspAlaLeuSerThrSerProPheHisLeuMetAsp 440
Db 1261 CGAACAGCAAAACCAACAGGTGACAGATGCTTGAAACACAGCCACCCACCTCATGTGAT 1320
Qy LeuLeuArgLeuHisSerSerProThrPheGlnValTyrGlyPheLeuArgAlaCysLeuCys 460
Db 1321 TTGCTCCGCTGACACAGACTCCCTGACAGTATATATGTTTCTTCGAGCCCTGTCTCC 1380
Qy LysValValSerAlaSerLeuThrPheArgHisAsnGlnArgArgPhePheLysAsn 480
Db 1381 AAGGTGTGTCTGTAGTCTCTGGGTAACAGGCAATATGAGCCGCGCTTCTTATAGAAC 1440
Qy LeuLysLysPheHisSerLeuGlyLysTyrGlyLysLeuSerLeuGlnLeuMetThr 500
Db 1441 TTAAAGAGTTCACTCTGTTGGGAAATACGGCAAGCTATCACTGACAGAACTGATGTGG 1500
Qy LysMetLysValGlnAspCysHisThrPheArgSerSerProGlyLysAspArgValPro 520
Db 1501 AAGATGAAGTAGAGATTGCCACTGGCTCCGACAGCCCGGGGAGAGACCGTGTCC 1560
Qy AlaAlaGlnHisArgLeuArgGlnArgGlyLeuAlaThrPheLeuPheThrLeuMetAsp 540
Db 1561 GCTGCAGAGACACCGTCTGAGAGAGAGATCTGGCTACGTTCTCTGCTGCTGATGAC 1620
Qy ThrTyrValValGlnLeuLeuArgSerPhePheTyrIleThrGlnSerThrPheGlnLys 560
Db 1621 ACATACCTGATCAAGCTGCTTATGCTTATTTTATCATCAAGAGACATTCACAGAG 1680
Qy AsnArgLeuPhePheTyrArgLysSerValTyrPheSerLysLeuGlnSerIleGlyValArg 580
Db 1681 AACAGGCTCTTCTCTACCTGTAAGAGTGTGTGAGAGACCTGACAGGCAATGGAATCGG 1740
Qy GlnHisLeuGlnArgValArgLeuArgGlnLeuSerGlnGlnValArgHisGln 600
Db 1741 CAACACCTTAAAGAGTGCAGGCTACGGAGCTGTCAACAGAGAGGTCAAGGATACCAAG 1800
Qy AspThrThrPheLeuAlaMetProIleCysArgLeuArgPheIleProLysProAsnGlyLeu 620
Db 1801 GACACCTGGCTAGCACTGCTCATCTGCAAGCTCGCTTCATCCCAAGCCCAACGGCTCG 1860
Qy ArgProIleValAsnMetSerTyrSerMetGlyThrArgAlaLeuGlyArgArgLysGln 640
Db 1861 CGGCCCATGTGAACATGAGTATAGCATGGGTACCAAGAGCTTTGGGCAAGAGAGAGAG 1920
Qy AlaGlnHisPheThrGlnArgLeuLysThrLeuPheSerMetLeuAsnTyrGlnArgThr 660
Db 1921 GCCAGCATTTCAACCAAGGCTCTCAAGACTCTTCAAGCATGTCTCAACTATGAGGGACA 1980
Qy LysHisProHisLeuMetGlySerSerValLeuGlyMetAsnAspIleTyrArgThrTrp 680
Db 1981 AATATGATCAAGCACTCGAGAGACACTACTGTATTCGCCAGATATGCAATGTGTCGGAGA 2040
Qy ArgAlaPheValLeuArgValArgAlaLeuAspGlnThrProArgMetTyrPheValLys 700
Db 2041 CGGSCCTTTGTGCTGCTGTGCTGCTGTGCAACCAACACCAAGATATGACTTTGTATAG 2100
Qy AlaAspValThrGlyAlaTyrAspAlaIleProGlnGlyLysLeuValGlnValAla 720
Db 2101 GCAGATGTGACCGGGGCTATGATGACCATCCCAAGGATAGGTGTGAGGTTGTTC 2160
Qy AsnMetIleArgHisSerGlnSerThrTyrCysAlaArgGlnTyrAlaValArgArg 740
Db 2161 AATATGATCAAGCACTCGAGAGACACTACTGTATTCGCCAGATATGCAATGTGTCGGAGA 2220
Qy AspSerGlnGlyGlnValHisLysSerPheArgArgGlnValThrThrLeuSerAspLeu 760
Db 2221 GATATGCCAAGGCCAAGTCCCAAGTCTTATGAGACAGAGTCAACCACTCTCTGACCTC 2280
Qy GlnProCtyrMetGlyGlnPheLeuLysHisLeuGlnAspSerAspAlaSerAlaLeuArg 780
Db 2281 CAGCATATAGATGGGCAGTTCCTTAAGCATCTGCAGATTCAGATTCAGATGCAGTACGAGG 2340
Qy AsnSerValValIleGlnGlnSerIleSerMetAsnGlnSerSerSerLeuPheAsp 800
Db 2341 AACTCGTGTATATGACAGAGCACTCTATATGATATGAGACAGACACACACCTGTGTGAC 2400
Qy PhePheLeuHisPheLeuArgHisSerValValLysIleGlyAspArgCysTyrThrGln 820
Db 2401 TTCTTCTGCACTTCTCTGTCACAGTGTCTTAAAGATGGTATGACAGTGTATAGCAG 2460
Qy CysGlnGlyIleProGlnGlySerSerLeuSerThrLeuLeuCysSerLeuCysPheGly 840
Db 2461 TGCAGGGCATCCCAAGGGCTCCAGGCTTATCCACCCTGCTCTGCACTGTGTGTGGA 2520
Qy AspMetGlnAsnLysLeuPheAlaGlnValGlnArgAspGlyLeuLeuLeuArgPheVal 860
Db 2521 GACATGAGAAACAACTGTTTGTCTAGATGCAGCGGATGGGTTGCTTTACGTTTGT 2580
Qy AspAspPheLeuLeuValThrProHisLeuAspGlnAlaLysThrPheLeuSerThrLeu 880
Db 2581 GATGACTTCTGTGTGTGACGCTCATTTGACCAAGCAAAACCTTCTCAGACACCTG 2640
Qy ValHisGlyValProGlnTyrGlyCysMetIleLeuLeuGlnLysThrValValAsnPhe 900
Db 2641 GTCCATGCGCTTCGATGATGGGTGCATGATTAACCTTGCAAGAGACAGTGTGAATCTC 2700
Qy ProValGlnProGlyThrLeuGlyGlyAlaAspProTyrGlnLeuProAlaHisCysLeu 920
Db 2701 CCGTGGAGACCTGATACCTGAGTGTGCATCTCATACAGCTGCGCTCATCTGCTG 2760
Qy PheProTyrCysGlyLeuLeuLeuAspThrGlnThrLeuGlnValPheCysAspTyrSer 940
Db 2761 TTTCCCTGTGTGGCTGTGTGTGTGACACTGACCTTGAGGGGTGTTCTGTGATTACTCA 2820
Qy GlyTyrAlaGlnThrSerIleLysThrSerLeuThrPheGlnSerValPheLysAlaGly 960
Db 2821 GGTATGCCCAACCTCATATTAAGCAAGCTCACTCCAGAGTGTCTCAAGCTCGG 2880
Qy LysThrMetArgAsnLysLeuLeuSerValLeuArgLeuLysCysHisGlyLeuPheLeu 980
Db 2881 AAGACCATCGGAAACAGACTCTGCTGCTGCTGCTGAGGTGACAGGCTATTTCTA 2940
Qy AspLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLysIlePheLeuLeu 1000
Db 2941 GACTTCAGATGAACAGCTCCACAGCTGTGCATCATATATACAGATCTTCTGCT 3000
Qy GlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArgValArgLys 1020
Db 3001 CAGGCTTACAGGTTCCATGATGATGTGATTCAGCTTCCCTTGACCAAGCGTATAGAGAG 3060
Qy AsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysTyrAlaIleLeu 1040
Db 1021 AACCTCAATCTCTTCTGAGCATCTCCAGCAAGCATCTGCTGTATGATATCTCTG 3120
Qy LysValLysAsnProGlyMetThrThrLeuLysAlaSerGlySerPheProGlnAlaAla 1060
Db 1041 AAGCTAAGAAATCCAGGAATGACATTAAGGCTGTGCTCTCTCCCTCGAAGCCGCA 3180
Qy HisThrLeuCysTyrGlnAlaPheLeuLeuLysLeuAlaHisSerValIleTyrLys 1080
Db 1061 CATTTGCTGTGCTTACCAAGCCCTTCTGCTCAAGCGTGTCTCATTTCTGTATCAAAA 3240
Qy CysLeuLeuGlyProLeuArgThrAlaGlnLysLeuLeuCysArgLysLeuProGlnAla 1100
Db 1081 TGTCTCTGAGACCTTGAAGACACCCCAAAACGTGTGTGCGGAAGCTCCAGAGGCG 3300
Qy ThrMetThrIleLeuLysAlaAlaAspProAlaLeuSerThrAspPheGlnThrIle 1120
Db 1101 ACATATACCATCTTAAAGTGCAGCTGACCCAGCCCTTAACACAGACTTTCAAGACATT 3360
Qy LeuAsp 1122
Db 1121 TTTGAC 3366
```



RESULT 2  
 AF051911 3426 bp mRNA linear ROD 01-DEC-2000  
 LOCUS  
 DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.  
 AF051911  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Scuriograthia; Muridae; Murinae; Mus.  
 REFERENCES  
 1 (bases 1 to 3426)  
 Greider, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.  
 Expression of mouse telomerase reverse transcriptase during  
 development, differentiation and proliferation  
 Oncogene 16 (13), 1723-1730 (1998)  
 MEDLINE  
 PUBMED 98241176  
 9582020  
 2 (bases 1 to 3426)  
 Greider, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.  
 Direct Substrates  
 Submitted (02-MAR-1998) Microbiology and Immunology, Albert  
 Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY  
 10461, USA  
 FEATURES  
 source  
 location/Qualifiers  
 1..3426  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="13"  
 /map=WGD offset position 40; 40 cm from centromere"  
 30..3398  
 /note="rTER"  
 /codon\_start=1  
 /product="telomerase reverse transcriptase"  
 /protein\_id="AAC00932.1"  
 /db\_xref="GI:3005592"  
 /translation="MTRAPCDPVRSLRSREYEWPLATFYRLGREGRLVPGDF  
 KYRTIYAGCIYCMHSGQPPDALSHQVSLKEIYARVORLCERRENTLAGGDF  
 LNEAGCGPMAFTSVRSYLPNTVITISVAGMMILSRGGDLVYLALHCAIY  
 LVPSGAYVCGSPYXICATTDIMPVSASRYRTPRVGRNTFNRLFLQIYSSROG  
 APPPLALSRGTRKLSLTSTSVSADKAKCYTPVREBERQVLPFSGKSVPS  
 ASRPEVTEAKKSLSSGKGVSDLSGVCCKHSSITLSPQNAQLRPFETRR  
 FYSRDQGERLNPFFLSNLPNLTGARRLVLEIFGSPRPSGLCRTHLSRRYV  
 OKRPFLOQLLVNHACQYVRLRSHCRFRANQVDALNTSPHMLDLRLHSDPW  
 VYGFLLACIKVYSASLMTGTHNRERPFKKLKEFIISKAGKSLQELMKKKVDCW  
 KSSSGGDRVAAASHRLRERLIATPLFMIDVYVOLLRSPFIESTFOGNRLFP  
 RSVSVKSGISGTRHLEVRVLRISQEVRHODMTLMPICRPFKPGKAPRIY  
 NMSYSNGTRALRRKQAHFTRLKTLSSMLNTERKTHPLNSSLKANDIYRWPR  
 FYRVAGVLDYPRMTFVADYTGADAIPOKLEVVANNTIRSESTCYROYAVRR  
 DSGGVGHKSRQVITSLDLPYVGQFLKQSDNALNSVVLKQISNMSSSSSI  
 PFPEFLFSLHVSVKIIGDCYTCOGGIPQGSLSITLCSLCFGDMENKLAERQDGL  
 LRVDVDFILVTHLDQATFLSTLVHVGVEGKINLQKVVPVPEPTLGAGAPLV  
 LPHGCLFVWCGELLDOTLEVDVSGVGAOTSIKSTFOSVYKAKTANGLSLV  
 LKHGCLFDLQVNSLQTCINIVKIFILGAVRPACTVCPDPPQRKTLPEFGITIS  
 SQASCCYALIKRYNCGNTLKASGSPPRALHMLCYQAFLLKLAHSVITYKLISLR  
 AQLCLRKLPKATMTLLAAADPASTDFOTILD"  
 BASE COUNT 746 a 979 c 907 g 794 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1, 67e-299 Length: 3426  
 Score: 5901.00 Matches: 1122  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-0432-460-2 (1-1122) x AF051911 (1-3426)  
 2Y 1 MethThArgAlaIrrAaGcYcSPcAlaIvaAaAgserLeuLeuAgsrAgyIrrAgGu 20

Db	30	ATGACCCGGGCTCTCGTTGGCCCGCGAGTGCAGTCTCTCTCTGCGACGCCGATACCCGGAG	89
OY	21	ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln	40
Db	90	GTTGTGGCCGCTGGCAACCTTTGTGTGGCGCTGTGGGCCGGAAGGAGCGGCGCTGTGGCA	149
OY	41	ProGlyAspProIlysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTP	60
Db	150	CCCGGGAGACCGAAGATCTACCGCAGCTTTGTGTGGCCAAATGCTTAGTGATGACATCGG	209
OY	61	GlySerGlnProProIcAlaAspLeuSerPheHisGlnValSerSerLeuIysGlnLeu	80
Db	210	GGGTCAACACCCCTCCACCCGCGACCTTTCCTTCACACAGTGCATCCCTTAAGAAGCTG	269
OY	81	ValAlaArgValValGlnArgLeuCysGlyArgAsnGluValArgAsnValLeuAlaPheGly	100
Db	270	GTTGCCAGGGGTTGTGGCAAGAGCTGTGGAGCCGCAAGAGAAGAAAGTGTGTGGCTTGGC	329
OY	101	PheGluLeuLeuAsnGluValArgGlyGlyProMetValPheThrSerSerValArg	120
Db	330	TTTGGAGTCTCTTAAAGAGGCGACAGGCGGGCTCCCATGGCTTCATCATAGAGGTGGGT	389
OY	121	SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValaTrpMetLeuLeu	140
Db	390	AGCTACTTGGCCAAACCTGTTATTGAGACCCCTCGTGCATGTGGTGATGATGCTACTG	449
OY	141	LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu	160
Db	450	TTGAGCCAGTGGGAGAGACCTGCTGTGTCTACCTGTGGCAACATGAGCTCTTATCTT	509
OY	161	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleValAla	180
Db	510	CTGTGGTCCCCCGAGTGTGGCTTACAGAGTGTGGGTCTCCCTGACCAATTGTGGCC	569
OY	181	ThrThrAspIleTPProSerValSerAlaSerTyrArgProThrArgProValGlyArg	200
Db	570	ACCAAGGAATGTGGCCCTCTGTGTCCGTACAGTTACAGGCCACCCGACCCGTGGGACGG	629
OY	201	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro	220
Db	630	AATTTCACTAACCTTAGGTTCTTTCACAAAGATCAAGACAGTAGTGGCCAGAGAACACCG	689
OY	221	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer	240
Db	690	AAACCCCTGACCTTGCATCTCGAGGTCAAAAGGACATCTAGTCCACCAATCAGT	749
OY	241	ValProSerAlaIysValArgCysTyrProValProArgValGluGluIyrHis	260
Db	750	GTCCTTTCAGCTAAAGAGGCGAGATGCTAATCTGTCCGACAGTGGAGGAGGACCCAC	809
OY	261	ArgGlnValLeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerPro	280
Db	810	AGGACAGGTCTACCAACCCCAACAGGAAATCAATGGGTGCAGATCTGTGTGGTCCCC	869
OY	281	GluValProThrAlaGluLysAspLeuSerSerLysGlyLysValSerAspLeuSerLeu	300
Db	870	GAGGTGCTACACGAGAAAGATTGTCTTAAAGAAAGGTGTCTGACCTGAGTCTC	929
OY	301	SerGlySerValCysCysLysHisLysProSerSerThrSerLeuLeuSerProProArg	320
Db	930	TCTGGGTGCGTGTGTGTAACAAACCCAGGTCCACATCTGTGTGCACACCCCGC	989
OY	321	GluAsnAlaPheGluLeuArgProPheIleGluThrArgHisPheLeuLysSerArgGly	340
Db	990	CAAAAGCTTTCACTCAGGCGCATTTATTGAGACAGACATTTCTTCACTCCAGGGGA	1049
OY	341	AspGlyGlnGluArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThr	360
Db	1050	GATGGCCAAAGACGCTTAAACCTCTCATCTCACTACAGCACTCCAGCTTAACTTACTG	1109
OY	361	GlyAlaArgArgLeuValGluIleAlaPheLeuGlySerArgProArgThrSerGlyPro	380
Db	1110	GGGCGCAGAGACTGTGTGAATCATCTTCTGTGGCTCAAGCCTTAGACATCAAGACCA	1169

QY 381 Leucylsargthrh:sargleuser:argarg:tyr:trp:glu:met:arg:pro:leu:phe:gin 400  
DB 1170 CTGTCAGGACACCCGCTATACGCTGATGAGTGGAGTGGGCGCCCTGTTCCACACG 1229  
QY 401 LeuLeuValAsnH:sal:agi:cy:gl:ntyr:val:arg:leu:leu:arg:ser:his:cy:arg:phe 420  
DB 1230 CTGCTGGTGAACATGACAGAGTGCATATATGTCAGACTCTCGAGCTGCATTCAGAGGTT 1289  
QY 421 ArgThrAlaAsn:ngln:ln:val:Thr:asp:Ala:leu:asn:Thr:ser:Pro:his:leu:met:asp 440  
DB 1290 CGAACAGCAACCAACAGGTGACAGATGCTTGAACACCGCCACCGCATTCATGAT 1349  
QY 441 LeuLeuArgLeuH:his:ser:ser:Pro:trp:gl:ln:val:tyr:gl:yl:phe:leu:arg:Ala:cy:leu:cy: 460  
DB 1350 TTGCTCCGCTGCACACAGATCCCTGGCAGGTATATGTTTCTTGCGGCTGTCTCTCC 1409  
QY 461 LysValValSer:Ala:ser:Leu:trp:gl:yl:Thr:Arg:His:ngln:Ala:Arg:Arg:Phe:leu:Asn 480  
DB 1410 AAGGTGGTGTCTGATGCTCTGGGGTACACAGCACATATGACCGCGCTTTTAAAGAC 1469  
QY 481 LeuLysLysPheH:leu:ser:leu:gl:yl:tyr:gl:yl:leu:ser:leu:ngln:leu:met:trp 500  
DB 1470 TTAAAGAGTTCACTCTGTTGGGAAATACGGCAAGCTATCATCTGCAGAACTGATGG 1529  
QY 501 LysMetLysVal:gl:u:asp:cy:his:trp:leu:arg:ser:ser:Pro:gl:yl:lys:asp:arg:val:Pro 520  
DB 1530 AAGATGAAGATGAGAGATTGCCACTGGCTCCGACAGCCCGGGAGAGACCGTGTCCCC 1589  
QY 521 AlaAla:gl:u:his:arg:leu:arg:gl:u:arg:Ala:leu:Ala:Thr:Phe:leu:phe:trp:leu:met:asp 540  
DB 1530 GCTGCAGAGACCCGCTGAGGAGAGAGATCTGCTACGTTCTGTTCTGCTGAGTGCAG 1649  
QY 541 ThrTyrValVal:gl:ln:leu:leu:arg:ser:Phe:phe:tyr:lle:Thr:gl:ln:ser:Thr:Phe:gl:ln:ys 560  
DB 1650 ACATACGTGTACACAGCTGCTTACGTTCTTTTACATCAACAGAGACATTCACGAAG 1709  
QY 551 AsnArgLeuPhePhe:tyr:arg:lys:ser:val:trp:ser:lys:leu:ngln:ser:lle:gl:yl:val:arg 580  
DB 1710 AACAGGCTCTTCTCTACCGTGAAGAGTGTGAGCAAGCTGCAGAGCATGTGAGTCAAG 1769  
QY 581 GluHis:leu:gl:u:arg:val:arg:leu:arg:gl:u:leu:ser:gl:ngln:gl:u:val:Arg:His:sgln 600  
DB 1770 CAACACCTTGAGAGAGTGGCTACGGAGCTGCACAAAGAGAGTCAAGCATCACAG 1829  
QY 601 AspThrTrpLeuAla:met:Pro:lle:cy:arg:leu:arg:Phe:lle:Pro:lys:Pro:asn:gl:yl:leu 620  
DB 1830 GACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCCAAGCCCAACGCGCTG 1889  
QY 621 ArgPro:lle:Val:Asn:Met:Ser:tyr:ser:Met:gl:yl:Thr:Arg:Ala:leu:gl:yl:Arg:arg:lys:gl:ln 640  
DB 1890 CGGCCCATTTGACACATGATTATAGATGGATGCCAGAGCTTTGGCAGAGAGAGACG 1949  
QY 641 Ala:gl:ln:His:phe:Thr:gl:ln:arg:leu:yl:Thr:leu:phe:ser:Met:leu:asn:tyr:gl:u:arg:Thr 660  
DB 1950 GCCCAGCATTTCAACCAAGCTCTCAAGACTCTTTCAGCATGCTCAACATATGACGGACA 2009  
QY 661 LysHis:Pro:His:leu:Met:gl:yl:ser:ser:Val:leu:gl:yl:Met:Asn:asp:lle:tyr:arg:Thr:trp 680  
DB 2010 AAACATCTCAACCTTATGGGTCTCTGTACTGGGTATGATATGACATTCACAGACCTGG 2069  
QY 681 ArgAla:Phe:Val:leu:arg:val:arg:Ala:leu:asp:gl:ln:Thr:Pro:Arg:Met:tyr:Phe:Val:lys 700  
DB 2070 CGGGCTTTGTGCTGCGTGGTGGCTCTGACACAGACACCCAGAGATGATCTTTGTTAAG 2129  
QY 701 Ala:asp:val:Thr:arg:val:tyr:asp:Ala:lle:Pro:gl:ngln:yl:lys:leu:val:gl:u:val:Ala: 720  
DB 2130 GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGGTAAAGCTGGTGGAGGTGTTGTTGC 2189  
QY 721 AsnMet:lle:Arg:His:ser:gl:u:ser:Thr:tyr:cy:his:leu:ngln:tyr:Ala:Val:Ala:arg: 740  
DB 2190 AATATGATCAGGCACTCGGAGACAGCATCTGTATCCGCCAGATATGACAGTGTCCGGAGA 2249

QY 741 AspSer:gl:yl:gl:ln:Val:His:lys:ser:Phe:arg:arg:ln:val:Thr:Thr:leu:ser:Asp:leu 760  
DB 2250 GATAGCAAGGCAAGTCCATGACAGTCTTTAGAGACAGGTCAACACCTCTGACCTC 2309  
QY 761 GluPro:tyr:Met:gl:yl:gl:ln:Phe:leu:yl:His:leu:gl:ln:Asp:Ser:Asp:Ala:leu:arg 780  
DB 2310 CAGCCATATACATGGGCGATGCTTATAGCATCTGCAGAGATGAGATGCCAGTGCAGTGG 2369  
QY 781 AsnSer:Val:Val:lle:gl:ln:Ins:er:lle:ser:Met:Asn:gl:u:ser:ser:ser:Leu:Phe:asp 800  
DB 2370 AACTCCGTTGATCATCAGACAGACATCTTATGATATGAGAGAGAGAGAGAGCTGTTAC 2429  
QY 801 Phe:Phe:leu:His:Phe:leu:arg:His:ser:Val:Val:lys:lle:gl:yl:Asp:Arg:cy:tyr:Thr:gl:ln 820  
DB 2430 TTCTTCCTGCACTTCTCGGTGCACAGTGTCCGTAAGATTTGGTGCAGAGTGTATACGAG 2489  
QY 821 Cys:gl:ln:gl:yl:lle:Pro:gl:ngln:yl:ser:ser:leu:ser:Thr:leu:leu:cy:ser:leu:cy:Phe:gl:yl 840  
DB 2490 TGCAGGGCATCCCCAGGAGCTCCAGCTTCCATCCCTGCTGTGAGTGTGTTGCGA 2549  
QY 841 AspMet:gl:u:Asn:lys:leu:Phe:Ala:gl:ln:val:gl:ln:arg:asp:gl:yl:leu:leu:Arg:Phe:Val 860  
DB 2550 GACATGAGAGAACAGCTGTTGTGTAGGTGACAGCGGATGGGTTCCTTTACGTTTGT 2609  
QY 861 AspAsp:Phe:leu:leu:Val:Thr:Pro:His:leu:asp:gl:ln:Ala:yl:ser:Thr:Phe:leu:ser:Thr:leu 880  
DB 2610 GATGACTTCTGTGTGTGAGGCTCACTTGACCAAGCAAAACCTTCTCAGACCTTG 2669  
QY 881 Val:His:gl:yl:val:Pro:gl:ln:tyr:gl:yl:cy:Met:lle:Asn:leu:ngln:ys:Thr:Val:Val:Asn:Phe 900  
DB 2670 GTCCATGGCTTCTCTAGTATGGGTGACAGATTAATCTGCAGAAACAGTGGTGAATTC 2729  
QY 901 Pro:Val:gl:ln:Pro:gl:yl:Thr:leu:gl:yl:Ala:Ala:Pro:tyr:gl:ln:leu:Pro:Ala:His:cy:leu 920  
DB 2730 CCGTGGAGGCTGTACCTCGGTGTGTGAGCTCCATACCACTCCCTGCTCACTGCTG 2789  
QY 921 Phe:Pro:tyr:cy:gl:yl:leu:leu:leu:Asp:Thr:gl:ln:Thr:leu:gl:ln:Val:Phe:cy:Asp:tyr:ser 940  
DB 2790 TTTCCTCGATGTGCTGCTGCTGCTGCAGCATCAAGCTTTGAGAGTGTCTGTGATCACTCA 2849  
QY 941 GlyTyrAla:gl:ln:Thr:ser:lle:tyr:Thr:ser:leu:Thr:Phe:gl:ln:ser:Val:Phe:yl:Ala:gl:yl 960  
DB 2850 GATTATCCCAAGCTCAATTAAAGAGACCTCACCTTCAAGAGTGTCTTCAAACTGGG 2909  
QY 961 Lys:Thr:Met:Arg:Asn:lys:leu:leu:ser:Val:leu:arg:leu:ys:cy:His:gl:yl:leu:Phe:leu 980  
DB 2910 AAGACCATGCGGAACAAGCTCGTGGTGTGCGGTGATGATGTCAGAGTCTATTTCTA 2969  
QY 981 Asp:leu:gl:ln:Val:Asn:ser:leu:ngln:Thr:Val:cy:lle:Asn:lle:tyr:lys:lle:Phe:leu:leu 1000  
DB 2970 GACTTGAGGTGAACAGCTCCACAGACTGCTGATCATATATACAAAGATCTTCTGCTT 3029  
QY 1001 Glu:Ala:tyr:Thr:Phe:His:Ala:cy:Val:lle:gl:ln:leu:Pro:Phe:asp:gl:ln:Arg:lys 1020  
DB 3030 CAGGCTTACAGGTTCATGATGATGTGATTCAGCTTCCCTTGACACAGGCTTATGGAG 3089  
QY 1021 Asn:leu:Thr:Phe:Phe:leu:gl:yl:lle:leu:ser:ser:gl:ln:Ala:ser:cy:tyr:tyr:Ala:leu:leu 1040  
DB 3090 AACCTCAGATTTCTTCTGGGCACATCTCCAGCAAGCATCTGCGTGCATATGCTATCTG 3149  
QY 1041 LysVal:lle:Asn:Pro:gl:yl:Met:Thr:leu:yl:Ala:ser:gl:yl:ser:Phe:Pro:Pro:gl:ln:Ala: 1060  
DB 3150 AAGGTCAAGATCAGAGATGACACATTAAGGGCTCGTCTCTTCTCCCTGAAAGCCCA 3209  
QY 1061 His:trp:leu:cy:tyr:gl:ln:Ala:Phe:leu:leu:yl:Ser:Val:Ala:His:ser:Val:lle:tyr:lys 1080  
DB 3210 CATGGCTGTCTACAGGCTTCTGCTCAAGCTGGCTGCTCATTTCTGTATCTACAA 3269  
QY 1081 Cys:leu:leu:gl:yl:Pro:leu:arg:Thr:Ala:gl:ln:lys:leu:leu:cy:arg:gl:yl:leu:Pro:gl:ln:Ala 1100  
DB 3270 TGTCTCTGGAGACTCTGAGAGACAGCCCAAAATGCTGTGCGGAAAGTCTCCCAAGGCG 3329  
QY 1101 Thr:Met:Thr:lle:leu:yl:Ala:Ala:Asp:Pro:Ala:leu:ser:Thr:Asp:Phe:gl:ln:Thr:lle 1120

```

DB      3330  ACAAATGACCATCTTAAAGCTGACGCTGACCCACCCCTTAAGCAAGACTTTCAGACCATT 3389
QY      1121  LeuASP 1122
DB      3390  TTGGAC 3395

RESULT 3
AF149012 4170 bp mRNA linear ROD 21-FEB-2001
LOCUS    Mesocricetus auratus telomerase catalytic subunit mRNA, complete
DEFINITION
ACCESSION AF149012
VERSION    AF149012.1 GI:6572639
KEYWORDS
SOURCE
ORGANISM  Mesocricetus auratus (golden Syrian hamster)
           Mesocricetus auratus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
           Mesocricetus.
REFERENCE 1 (bases 1 to 4170)
AUTHORS  Guo, M., Okamoto, M., Lee, Y. M., Baluda, M. A. and Park, N. H.
TITLE     Enhanced activity of cloned hamster TERT gene promoter in
           transformed cells
JOURNAL   Biochim. Biophys. Acta 1517 (3), 398-409 (2001)
MEDLINE   21240330
PUBMED    11342218
REFERENCE 2 (bases 1 to 4170)
AUTHORS  Guo, M., Okamoto, M., and Park, N. H.
TITLE     Direct Submission
SUBMITTED (06-MAY-1999) Dentistry, University of California, Los
           Angeles, 10833 Le Conte Ave., Los Angeles, CA 90095, USA
FEATURES
     source
         1..4170
             /organism="Mesocricetus auratus"
             /mol_type="mRNA"
             /db_xref="taxon:10036"
             /note="common: golden Syrian hamster"
             28..3414
                 /note="hamTERT"
                 /product="telomerase catalytic subunit"
                 /protein_id="AA17334.1"
                 /db_xref="GI:6572640"
                 /translation="MPRAVRCRAVALLISQRYQVWPLATFVRLGEGSQLVQVQDP
                 KVFRTIARCLVCPMDSQPPADLSFQVSLKVARVQRLCEGSENVATFGA
                 LINGAGGCPMTFTSVRSYLPNSVTSRLVSGAMMLINLVQDGLVYLALCALYL
                 LVPSGAYOVGSPLYQICATAETWSPYSRIYPRTPYGRNFTLGSFVTEHVNLSHOB
                 AKKPPPLREAKRSLSITNRSVSPSKKACDIAPREKGPYQAVTPBDKTVWPP
                 AKSHAVPIRTTKEDLSGCVKAPGLSRGSGVCKEKPSTSLQSLQNAFOLAPYTE
                 TRKPLYSREGRELRNLPSFLNNLQPSLTGARRVETILFLMRRTSGPICGRRLRSK
                 RYQWMPPLFQQLVNHARCVPYRLRSLRCRFTAAHQVAGALNTSPQRLMRLRHS
                 SPQVQVFLQACVGLVPPGLGSGRHHNRREFKVKERISLGKDKLSLQELTWKXV
                 QDCRMRLSSPGNNCVPAAEHRTREIRILAVFLMLDAVYVLLRSPVYETFKMAY
                 LEFYKSMWRLOSIGVRLHREVRLOELSGEVRQEAHPAMPICRLAIFIPKPSGL
                 RPIYVNSVWGCTAPFGKQAOHFTQCLTSEVNLVELTKTNLGLASVGLNDIYT
                 WRRTVRYRTLDPAKMYFVADYTGAYDAIIPQDKLVITANMTRHNDNSGCHQVAV
                 VQDRQGLHKFSKFRQVSTLSDLPJHMGQDLKHODSTSLNLSVETESLSNBS
                 SSLPDLFLFVRSNVVKGRCYVQCGIPQSSLSLTLCSLCEGDMENKLPFVQDD
                 GLILRFVDDFLVTPHLVQAEAFRLAVRGIPKYMNILOKTVNPEVPAVAGTIDGA
                 PHQLPAHCFPMWCGGLDQTLVLCADYTGARTSIKASLTFORTFAGRMRSKLLA
                 VLAKCHSLFLDKNNSLQTCINVKYKIFLOAYFHAQLQLPDQVGRNRPFFLS
                 ITSNKSCCYSTLKNNAGMTKAGAGSGSPPEAARWLYQAFLLKLAGHSVYKCL
                 LGRFTRQKQLCRKLPRAIMATLRTAADPALSTDPQTLID"
BASE COUNT  927 a 1177 c 1126 g 940 t
ORIGIN

```

## Alignment Scores:

Score: 4751.00  
Percent Similarity: 88.04%  
Best Local Similarity: 80.07%  
Query Match: 80.51%

3.15e-239

Length: 4170

Matches: 904

Conservative: 90

Mismatches: 127

Indels: 8

```

DB:      10      Gaps:      4
US-09-042-460-2 (1-1122) x AF149012 (1-4170)
QY      1  MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
DB      28  ATGCCCGCCGCCGCCGTTGCGCGGCGCGCGCGCTGCTGCTGCGGACATACCGGTAG 87
QY      21  ValThrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
DB      88  GTGTGGCGCGCTGGCAACCTTCGTGCGCGCGCTGGCACTCGAGGAGGACGACCTGTAC 147
QY      41  ProGlyAspProIysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
DB      148  CCGCGGGACCAAGAGCTTCGCGACGTTGTGGCGCGCGCGCTGCTGCTGCTGCTGCTG 207
QY      61  GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGlyLeu 80
DB      208  GACTCACACCTCCACTGCTGACCTTCTTCACACAGGTGTCATCTGACGAGACTG 267
QY      81  ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
DB      268  GTGCCGACGGCTCGCGACAGACTCTGCGAGCGCGGAGAGGAGAGAGTGTGACTTTGGC 327
QY      101  PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
DB      328  TTCCGCGCTGCTTAAAGGAGCGGCGGCTCCCATGACATTCACACAGCGTGGC 387
QY      121  SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaThrMetLeu 140
DB      388  AGCTACCTGCCCACTGCGAGCTGAGTCTCTGCGGTAGTGTGCTTGATGCTTGG 447
QY      141  LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
DB      448  CTGAACCGAGTGGCGGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507
QY      161  LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
DB      508  CTGGTGGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 567
QY      181  ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
DB      568  ACCGAGAAACCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
QY      201  AsnPheThrAsnLeuArgPheLeuGlnGlnIleLeuSerSerSerArgGlnGluAlaPro 220
DB      628  AATTTTACTCATCTTGATGATCCACACCGGCTGAGACGAGCTCACCGAGAGCATGG 687
QY      221  LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
DB      688  AAACCCCGCGCTTGCATCTGATCCATCTCGAGAGCGGAGGAGTCTAGCATCACCAATAG 747
QY      241  ValProSerAlaValGlyAlaArgCysTyrProValProArgValGluGluGlyProHis 260
DB      748  GTGCTTCATCTTAAGAGAGCGGAGTGTGCTGCGCGGAGCTGAGAGAGGAGACCTTAC 807
QY      261  ArgGlnValLeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerPro 280
DB      808  AGCGAGGAGATTCACACCCATCAGACAAACATGAGGTGCACAAATCTGCAAGTCCCAT 867
QY      281  GluValPro-----ThrAlaGluLysAspLeuSerLysGlyValSerAsp 297
DB      868  CGAGGCGCTATTAGTACCACTACCAAGAGATTTGCTTCGCGGAGTGAAGGACCTGGC 927
QY      298  LeuSerLeuSerGlySerValCysCysLysHisLysProSerSerThrSerLeuLeuSer 317
DB      928  CTGAGTGGCTCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
QY      318  ProProArgGlnAsnAlaPheGlnLeuArgProPheIleGluThrArgHisPheLeuTyr 337
DB      988  CCACGTGTGCAAAATGCTTCACTGACCTACAGACATATACTAGACCAAACTCTTCTAC 1047
QY      338  SerArgGlyAspGlyGlnGluArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnPro 357

```

```

Db      1048  TCTAGGAGAGGTGGCCGAGAGAGGCTGAACCCCTCGTTCCTACTCAACCACTGCAGGCC 1107
Qy      358  AsnLeuThrGlyAlaArgArgLeuValGluIlePheLeuGlySerArgProArgThr 377
Db      1108  AGCTTGACTGGGGCCAGAGAGACTGTGAGAGACTCTTTCTAGGCACTGAGACTTAGACA 1167
Qy      378  SerGlyProLeuCyArgThrHisArgLeuSerArgThrTyrTrpGlnMetArgProLeu 397
Db      1168  TCGGACCACTCTGTGGAGAGCCGCCCTATCGAAGGCTACTGGCAGATGGCGCCCTTA 1227
Qy      398  PheGlnGlnLeuLeuValAsnHisAlaGlnCysGlnTyrValArgLeuLeuArgSerHis 417
Db      1228  TTCCAGCAGTTGCTGTGGAACCATGCCGGGCGCCGATGTCCGGCTCCAGTCCCAT 1287
Qy      418  CysArgPheArgThrAlaAsnGlnGlnValThrAspAlaLeuAsn---ThrSerProPro 436
Db      1288  TGCAGGTTTCGACCGGAGCCCAACAGGTGGCAGTGCCTTGAACACCAACCAACCCACAG 1347
Qy      437  HisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyrGlyPheLeuArg 456
Db      1348  CGCCCTCATGAATTGCTCCGTCTACAGACAGAGTCCCTGGCAGGATATGGCTTCTTCAG 1407
Qy      457  AlaCysLeuCyArgValValSerAlaSerLeuTrpGlyThrArgHisAsnGlnArgArg 476
Db      1408  GCCTGTGTCCGAAAGCTGTGCTCCGGGTCTCTGGGGTTCCCGGCAACAACCGCAGCC 1467
Qy      477  PhePheLeuAsnLeuLeuValPheHisSerLeuGlyLysTyrGlyLysLeuSerLeuGln 496
Db      1468  TTCTTTAAGAACGTGAACGGTCTATCTCTGGGGAAGTATGACAAAGCTGTGGCTCGAG 1527
Qy      497  GlnLeuMetTrpLysMetLysValGlnAspCysHisTrpLeuArgSerSerProGlyLys 516
Db      1528  GAGCTGACGTGGAGATGAAGTCAAGACTGCAGTGGCTGCAGCAACCCAGGGAAC 1587
Qy      517  AspArgValProAlaAlaGlnHisArgLeuArgGlnArgIleLeuAlaThrPheLeuPhe 536
Db      1588  AACTGTGTCCGGCTGCAGAGACCGGCAAGGGAAGAGATCCGTGTGTGTGTGTC 1647
Qy      537  TrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyrIleThrGlnLys 556
Db      1648  TGGCTGATGGACCGCTAGCTGTGAGCTGTCTTGCTCATTTCTTTACGTACAGAGAAC 1707
Qy      557  ThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSerLysLeuGlnSer 576
Db      1708  ACTTTCAGAAAGAACCGGCTCTCTTCAACGAAAGAGCATGTGGAGAAAGCTGCAGAGC 1767
Qy      577  IleGlyValArgGlnHisLeuGlnArgValArgPheArgGlnLeuSerGlnGlnGlnVal 596
Db      1768  ATTGAGTCAAGSACACACTTGAGAGAGTGGGCTACAGAACTGTCTCAAGAAAGAGTC 1827
Qy      597  ArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArgPheIleProLys 616
Db      1828  AGGCAGCGCCGAGAGAGCCCTGGCAGGCCATGCCCATCTGCAGACGCTTTTCATCCCAAG 1887
Qy      617  ProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThrArgAlaLeuGly 636
Db      1888  CCCAGTGGCTTCCGCCCATTTGTGAACATGAGTTAT---ATGGCACAACAGCTTTGAC 1944
Qy      637  ArgArgLysGlnAlaGlnHisPheThrGlnArgLeuArgLysThrLeuPheSerMetLeuAsn 656
Db      1945  AAAGGGAAGCAGAGCTCAGCATTTCCACAGTGTCTCAAGACTCTGTTCAACGCTGCACAC 2004
Qy      657  TyrGlnArgThrLysHisProHisLeuMetGlySerSerValLeuGlnMetAsnAspIle 676
Db      2005  TATGAACCTGACAAACATACCTACTGGGGGAGATGTGACTGGGCTCGAATGATATC 2064
Qy      677  TyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGlnThrProArgMet 696
Db      2065  TACAGGACCTGGCGAGCTTGCTACTGGGTGAGGCACTCGAACCAAGACCAAGAGATG 2124
Qy      697  TyrPheValLysAlaAspValThrGlyAlaTyrPheAlaIleProGlnGlnLysVal 716

```

---

```

Db      2125  TACTTTGTAAAGCAGATGTGACAGGGGATATGATGTCATCCCCCAGAGACAAGTTGTG 2184
Qy      717  GlnValValAlaAsnMetIleArgHisSerGlnSerThrTyrCysIleArgGlyThrAla 736
Db      2185  GAGGTATTTGCCAATATGATGATCAACACCAACCACTGCTATGATCAACCAATATCA 2244
Qy      737  ValValArgArgAspSerGlnGlnValHisLysSerPheArgArgGlnValThrThr 756
Db      2245  GTGGTCAAGAAATATGACAAAGCCAAATCCAAAGTCTTCAAGAGACAGGTTCCAC 2304
Qy      757  LeuSerAspLeuGlnProGlyMetGlyGlnPheLeuValHisLeuGlnAspSerAla 776
Db      2305  CTCCTGACCTCCAGGCACACAGGGCCAGTTCTTGAAACATCTTCAGGACTCAACACC 2364
Qy      777  SerAlaLeuArgAsnSerValValIleGlnGlnSerIleSerMetAsnGlnSerSer 796
Db      2365  AGTCGGTGAAGAACTCCGTGTGATGACAGAGTTTCTCTGAACAGAGCCAGCAGCC 2424
Qy      797  SerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLysIleGlyAspArg 816
Db      2425  AGCTGTTCGACTTCTTCCGCTTGGCGTATGACAGTGTGTAAGATCCGTGGCAGG 2484
Qy      817  CysThrThrGlnCysGlnGlyIleProGlnGlnLysSerLeuSerThrLeuLeuCysSer 836
Db      2485  TGTATGTCCAGTGTCCAGGCATCCCAAGGGTCCAGGCTGTCCACCTGTCTGCACT 2544
Qy      837  LeuCyAspPheGlyAspMetGlnAsnLysLeuPheAlaGlnValAlaArgAspGlyLeuLeu 856
Db      2545  CTGTGTTTCGGGACATAGAGAAACAGTGTGTGCTGAAGTCAACAGAGATGGCTGCTT 2604
Qy      857  LeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGlnAlaLysThrPhe 876
Db      2605  TTGGTTTGTATGACTTCTGTGTGGAGAACCTCAACTGTGTCCAGGCGGAACCTTC 2664
Qy      877  LeuSerThrLeuValHisGlyValProGlnTyrGlyCysMetIleAsnLeuGlnLysThr 896
Db      2665  CTCAGGACCTCTCCGTGGCAACCTTGAGTACGGGTGATATTAACCTGCAGAAAGACA 2724
Qy      897  ValValAsnPheProValGlnProGlyThrLeuGlyValAlaAspProGlyGlnLeuPro 916
Db      2725  GTGGTAACCTTCCCTGGAGCGCTGTATACCTGATAGGACAGCTCCACACAGCTGCT 2784
Qy      917  AlaHisCysLeuPheProTrpCysGlyLysLeuLeuLeuAspThrGlnThrLeuGlnValPhe 936
Db      2785  GCTACGTGCTGTTCCTGTGTGGCTTACGTGCTGACATCACTGATCAATGAGTGTCTC 2844
Qy      937  CysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThrPheGlnSerVal 956
Db      2845  TGTGACTACCTGGTTATGCGCGACCTCAATTAAAGCCAGCTCACTTCAGCGCACCC 2904
Qy      957  PheValAlaGlySerThrMetArgAsnLysLeuLeuSerValLeuArgLeuLysCysHis 976
Db      2905  TTCAAGCGGGAGAGAACTGACACAGAAAGCTTTCAGCTGTGTGGGTGAGTGCAC 2964
Qy      977  GlnLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLys 996
Db      2965  AGTCTGTTTCTAGACTTGACAGATGATAGCTTCAACAAGTCTGATCAATGATGACAG 3024
Qy      997  IlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGln 1016
Db      3025  ATCTTCCTGTCTCAGGCTCAAGCTTCAGATTCAGCTGTGGCTCAAGTCTTTCAGCA 3084
Qy      1017  ArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysCys 1036
Db      3085  CATGTTGAAGAAAGAACCCGCAATCTTCTGAGATCATCTCAACATACCACTGCTGCC 3144
Qy      1037  TyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLys-----AlaSerGly 1053
Db      3145  TACTCCATCTCGAAGGTCAGAAATGACAGATACACTAAAGGCCAAAGGCTCCTGCGC 3204
Qy      1054  SerPheProProGlnAlaAlaHisTrpLysCysTyrGlnAlaPheLeuLysValAla 1073
Db      3205  TCAATTCCTCTGAAGCTGACAGTGGCTCTGTTACAAAGCTTCTGCTCAAGCTGGCT 3264

```



Db 1216 GGGGNGCTCTCAAGAGCACTGCCGCTGGAGCTGCG-----GTCACCCCGACGA 1266  
Qy 431 Leu-----Aan 432  
Db 1267 GCCGGTCTGTGCCCCGAGAGACCCGAGGCTGTGTGGCGGCCCCCGAGAGAGAGAC 1326  
Qy 433 ThSerProProHISLeuMetAspLeuLeuArgLeuHISerSerProTropJinValYyr 452  
Db 1327 ACAGACCCCGCTGGCTGTGAGCTCTCCGACAGACAGAGCCCTGGCAGCGTGC 1386  
Qy 453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTropJinYThrArgHIS 472  
Db 1387 GCGTTCGTGCGGCGCTGTGCTGCGCGCTGGTCCCGAGCCCTCTGGGCTCCAGGAC 1446  
Qy 473 AengLuarArgPhePheLeuAsnLeuLeuYlsPheHISerLeuGlyYyrGlyYls 492  
Db 1447 AACGAAAGCCGCTTCTCTAGAGAACCCAGAAAGTTCACTCTCTGGGAGAGCATGCCAAG 1506  
Qy 493 LeuSerLeuGlnGluLeuMetTropLysMetLysValGluAspCysHISITPLeuArgSer 512  
Db 1507 CTCTGCTGCAGAGAGCTGACGTGGAAGATGAGCGTGGGAGCTGGCTGGCCAGG 1566  
Qy 513 SerProGlyYyrAspArgValProAlaAlaGlnHISArgLeuArgGluArgIleLeuAla 532  
Db 1567 AGCCAGGGGTTGGCTGTGTCCGCGCAGAGACCGTGTGGTGAAGAGATCTGACC 1626  
Qy 533 ThrPheLeuPheTropLeuMetAspThrYyrValValGlnLeuLeuArgSerPhePheYyr 552  
Db 1627 AAGTCTCGACCTGGCTGATGAGTGTGAGTGTGAGCTGCAGCTCTTTCTTTAT 1686  
Qy 553 IleThrGlnSerTropPheGlnLysAsnArgLeuPhePheYyrArgLysSerValITPser 572  
Db 1687 GTACAGAGACACAGCTTCAAAAGAGCGCTCTTTTCAACCGAAGAGTGTCTGGAGC 1746  
Qy 573 LysLeuGlnSerIleGlyValArgGlnHISLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1747 AAGTTCAAAGCATTTGAAATGACACACACTTGAAGAGGTGACGCTGGGAGAGCTGTGC 1806  
Qy 593 GlnGluGluValArgHISHisGlnAspThrTropLeuLysMetProIleCysArgLeuArg 612  
Db 1807 GAAGCAGAGGTGCAGGCGCATGCGGAAGCCAGGCGCCCTGCTGACGTCACAGACTGCC 1866  
Qy 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerYyrSerMetGlyThr 632  
Db 1867 TTCACTCCCAAGCTGACGGGCTGGCGCATGTGTGAACATGACATGCGTGGGAGGCC 1926  
Qy 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHISerThrGlnArgLeuYsrThrLeuPhe 652  
Db 1927 AGAAGTTCGCGAGAAAGAGAGGCGGAGCGTCTCACTGAGGGTGAAGGACCTGTTTC 1986  
Qy 653 SerMetLeuAsnTyrGluArgThrLysHISProHISLeuMetGlySerSerValLeuGly 672  
Db 1987 AGCGTGCACACTACAGAGCGGCGCGCGCGCGCGCTCTGGAGCGCTGTGTGGCGGC 2046  
Qy 673 MetAsnAspIleTyrArgThrTropArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2047 CTGACGATATCCACAGGCGCTGGCGACCTGTGCTGTGTGGGGGCCACAGACCG 2106  
Qy 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValArgAlaIleProGln 712  
Db 2107 CCGCTGAGCTGTACTTGTTCAGAGGTGATGTGACGGGCGCGTACACACATCCGCCAG 2166  
Qy 713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHISerGlnSerTropJinCysIle 732  
Db 2167 GACAGGCTCAGGAGGTATATGCCAGCATATCAAA---CCCGAAGACAGTACTGCGG 2223  
Qy 733 ArgGlnTyrAlaValAlaArgArgAspSerGlnGlyGlnValHISLysSerPheArgArg 752  
Db 2224 CGTGGTATGCCGCGGTCTCAGAGAGCGCGCCATGGGACCTCCGCAAGGCGCTTCAAGAGC 2283  
Qy 753 GlnValThrThrLysSerAspLeuGlnProYyrMetGlyGlnPheLeuYsrHISLeuGln 772  
Db 2284 CACGCTCTACCTTGAACAGACTCCAGCCGTATATGCGAAGTTCTGTGGCTCACTGACG 2343

Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGlnGlnSerIleSerMetAsn 792  
Db 2344 GAGACC-----AGCCCGCTGAGAGATGCGCTGTATCTGACACAGAGCTCTCTGAAAT 2397  
Qy 793 GluSerSerSerSerLeuPheAspPhePheLeuHISPheLeuArgHISerValValLys 812  
Db 2398 GAGCGCAGCAGTGGCTCTTCAACGCTTCACTACGCTTATGTCACACAGCGGTGCG 2457  
Qy 813 IleGlyAspArgCysYyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
Db 2458 ATCAGGGCGAAGTCTTACGTCAGTGCAGGGGATCCCGAGGCGCTTCACTCTTCCACG 2517  
Qy 833 LeuLeuCysSerLeuLeuPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
Db 2518 CTGCTCTGACGCTGTGTGATACGCGCATGAGAACAGACTGTTGGGGGATTTGGGCG 2577  
Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHISLeuAspGln 872  
Db 2578 GACGGGCTGCTCTGGCGTTGGTGAATGATTTCTGTGTGACACCTCACCTCACCCAC 2637  
Qy 873 AlaLeuThrPheLeuSerThrLeuValHISGlyValProGluTyrGlyCysMetIleAsn 892  
Db 2638 GCGAAACCTTCTCTCAGACCTGATCCAGAGTGTCTCACTATGCTGCTGGTGAAC 2697  
Qy 893 LeuGlnLysThrValValAsnProValGluProGlyThrLeuGlyValAlaIlePro 912  
Db 2698 TTGGGAAAGACATGTGTGAATCTCTCTGTAAGAAAGAGAGCCCTGGGTGGCAGGCTTTT 2757  
Qy 913 TyrGlnLeuProAlaHISCysLeuPheProTropYyrGlyLeuLeuLysAspThrGlnThr 932  
Db 2758 GTTACATGCGCGGCCCAAGCTATTCCTGTGGCGCTGTGCTGTGTGATACCGAGACC 2817  
Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2818 CTGAGGCTCAGAGCGACATCTACACTATGCGCGACCTCATCAGAGCGAGTCTCAC 2877  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2878 TTCAACCGCGCTTCAAGGTGGAGAAACATGCGTGGAACTTTTGGGGTCTTGGCG 2937  
Qy 973 LeuLysCysHISGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2938 CTGAAGTTCACAGCTGTCTTCTGATTTGACGTGACAGGCTTCCAGAGCGTGTGCAC 2997  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHISAlaCysValIleGlnLeu 1012  
Db 2998 AACATCTAACAAATCTCTGCTGCAGGCGTACAGGTTCAAGCATGTGTGTGCAGCTC 3057  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3058 CCATTTATACAGAAAGTTGAAAGAACCCCAACATTTTCTCGGCTCAATCTTACACG 3117  
Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValIysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3118 GCGTCCCTGTGATCTACCTCGTGAAGGCAAGAACGAGGATGCTGTGGGCGCAAG 3177  
Qy 1053 GlySer-----PheProProGluAlaAlaHISITPLeuCysTyrGlnAlaPheLeu 1069  
Db 3178 GGGCGCGCGCCCTCTGCTCCGCGGCGGTGCGAGTGTGTGACACCAAGATCTCTG 3237  
Qy 1070 LeuLysLeuAlaAlaHISerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3238 CTCAAGTGACTGACACCGTGTCACTTACGAGGCACTCTGGGTCACTCAGAGACACG 3297  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3298 CAGACGAGCTGTGTGTGAAGCTCCCGGAGACAGAGCTGACCTGAGAGCGCGACGC 3357  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3358 AACCCGCACTGCGCTCAGACTTCAAGACCATCTGTGAC 3396



RESULT 5  
 BD094749  
 LOCUS  
 DEFINITION  
 The cell having the potentiality of differentiation into  
 cardiomyocytes.  
 ACCESSION  
 BD094749  
 VERSION  
 BD094749.1 GI:22640337  
 KEYWORDS  
 MO 0148150-A/16  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (bases 1 to 3396)  
 Umezawa, A., Hata, J., Fukuda, K., Ogawa, S., Sakurada, K., Gojo, S. and Yamada, Y.  
 The cell having the potentiality of differentiation into  
 JOURNAL  
 Patent: WO 0148150-A 16 05-JUL-2001,  
 KYOWA HAKKO KOGYO CO. LTD, AKIHITO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, KAZUHIRO SAKURADA, SATOSHI GOJO, YOUTI YAMADA  
 COMMENT  
 OS Homo sapiens (human)  
 PN WO 0148150-A/16  
 PD 05-JUL-2001  
 PF 02-NOV-2000 WO 2000JP007741  
 PR 28-DEC-1999 JP 99P 372826, 28-FEB-2000 WO PCTJP0001448 PI  
 AKIHITO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, PI  
 KAZUHIRO SAKURADA, SATOSHI GOJO, YOUTI YAMADA  
 PC C12N5/06, C12N5/10, C12N15/09, A61K31/203, A61K35/28, A61K38/19, PC  
 A61K38/39,  
 PC A61K38/45, A61K48/00, A61P9/10, A61P41/00, C07K16/28, C12P21/08, PC  
 C12Q1/02,  
 PC C12Q1/48, G01N33/577  
 CC The cell having the potentiality of differentiation into CC  
 FH Key Location/Qualifiers  
 FT CDS (1)..(3399).  
 Location/Qualifiers  
 1..3396  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 549 a 1157 c 1088 g 602 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,62e-174 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: Gaps: 13  
 US-09-042-460-2 (1-1122) x BD094749 (1-3396)  
 QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 DB 1 ATGGCGGCGCGCTCCCGCTCGCGACCGCGGCGCTCGCTCGCGACGACACACCGCGAG 60  
 QY 21 ValTTrProLeuAlaTrpPheValArgSerGluGlyProGluGlyArgArgLeuValGln 40  
 DB 61 GTGCTGCGGCTGGCGACGTTGTCGGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCTGGTGA 120  
 QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 DB 121 CGCGGGGACCGCGGCGCTTCGCGCGCGCTGGTGGCGCGCGCTGGTGGTGGCGCGCG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGlnLeu 80  
 DB 181 GACGACG 240  
 QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 DB 241 GTGCGCGGAGTGTGCGAGAGGCTGTGGAGCGCGCGCGCGAGAAAGTGTGTGGCGCTTGGCG 300

QY 101 PheGluLeuLeuAsnGluValArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 DB 301 TTGGCGGTGTGAGACGGGGCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 360  
 QY 121 SerTyrLeuProAlaThrValIleGluThrLeuArgValSerGlyValTrpMetLeuLeu 140  
 DB 361 AGTAACTGTCCCAACAGAGTGAACGACGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420  
 QY 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 DB 421 CTGGCG 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 DB 481 CTGGTGGCTCCACGCTCGCTGACAGTGTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCT 540  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 DB 541 GCCACTGACGGCG 594  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220  
 DB 595 -----TGCAACGGCGCTGGAACCATGCGTCAGGAGCGCGGG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 DB 634 GTCCCGCTGGCGCTGCGAGCGCGGGTGCAGAGAGCGGGGGCGAGTGCACCGCAAGT 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGlnGlnGlyPro--- 259  
 DB 694 CTGGCGGTGCCAAGAGGGCCAGCGTGGCGTGGCGCTGACCGCGAGCGAGCGCGCTT 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 DB 754 GCGGAGGGGTCTGGCGCCACCGCGAGAGCGGTGACCGAGTACCGTGTCTGTCTGT 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
 DB 814 GTGGTGTACCTGGCGAG-----CCGCGCAAGAGCACCTCTTGGAG 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 859 GGTGGCTCTCTGGCG 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 919 CCCCATTCAATCG 978  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 979 GAGACCAAGCACTTCTCTACTCTCTCAGCGCAC--AAGAGCAGCGTGGCGCGCTCTTC 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 DB 1036 CTACTGAGTCTCTAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisAspLeuSerArg 390  
 DB 1096 CTGGGTTCAGCGCGCTGATCCAGAGCATCCCGCGAGGTTGCCCGCGCGCGCGCGCG 1155  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnIleLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1156 TACTGCAATAGCG 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 DB 1216 GGGGTCTCTCAAGACGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
 QY 431 Leu-----Asn 452  
 DB 1267 GCGGTGTGTGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1326  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgHisSerSerProTrpGlnValTyr 452

Db 1327 ACAGACCCCGCTCCCTGGTGGACAGTCTGCTCCGACAGACAGCCCTCGGACAGGTATC 1386  
 Qy 453 GlyPheLeuArgAlaCyLeuCyLeuValValSerAlaSerLeuTProLysThrArgHis 472  
 Db 1387 GGGCTTCGTGGGGCTCCCTCCGAGACACCAAGATTCTCTCCGAGGAGATCTCCAGGAC 1446  
 Qy 473 AsnGluArgArgPhePheLeuValLeuValPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 492  
 Db 1447 AACAGACGCGCTTCCTCCAGACACCAAGATTCTCTCCGAGGAGATCTCCAGGAC 1506  
 Qy 493 LeuSerLeuGlnGluLeuMetLysValGluAspCysHisSTriLeuArgSer 512  
 Db 1507 CTCTCGCTGCAGAGCTGACGTGGAGATGAGCGTGGAGCTGCCCTTGGCTGCAGG 1566  
 Qy 513 SerProGluLysAspArgValProAlaAlaGluHisArgPheArgGluArgGluLeuAla 532  
 Db 1567 AGCCAGAGGAGTGGCTGTGTTCCGCGCAGAGACCCGTCGTGAGAGAGATCTGAGCC 1626  
 Qy 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 Db 1627 AAGTTCTGCACCTGCTGATGATGTGATGATGATGATGATGATGATGATGATGATGAT 1686  
 Qy 553 IleThrGluSerTrpPheGlnLysAsnArgLeuPhePheTyrArgLysSerValITPser 572  
 Db 1687 GTCAAGGAGACCAAGTTCACAAAGACAGCTCTTTTCTACCGGAGAGGTGTCTGGAGC 1746  
 Qy 573 LysLeuGlnSerIleGluValArgGlnHisIleGluGluArgValArgLeuArgGluLeuSer 592  
 Db 1747 AAGTTGCAAAAGCATTTGAATGACACACACTTGAAGAGGTGTGAGCTGGGAGAGCTGTG 1806  
 Qy 593 GlnGluGluValArgHisIleGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
 Db 1807 GAAGCAGAGGTCAAGCAGCATCGGAGAACAGCCGCTGTGACGTCCACAGCTCCAGC 1866  
 Qy 613 PheIleProLysProAsnGluLysArgProIleValAsnMetSerTyrSerMetGlyThr 632  
 Db 1867 TTCAATCCCAAGCTCGACGGGCTCGGCGCATTTGTGAACATGACATACGTCGTGGAGAC 1926  
 Qy 633 ArgAlaLeuGluArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
 Db 1927 AGAACGTTCCGACAGAAAGAGGCGCAGCGCTCACTCGAGGGTGAAGGACAGCTGTC 1986  
 Qy 653 SerMetLeuAsnTyrGluArgThrLysHisProHisIleuMetGlySerSerValIleuGly 672  
 Db 1987 AGCGTGTCAACTACAGAGCGGCGCGCGCGCGCTCCGCGCGCTGTGTGTGTGTGTGTGT 2046  
 Qy 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
 Db 2047 CTGACGATATCCACAGAGGCGCTGGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2106  
 Qy 693 ThrProArgMetTyrPheValLysAlaAspValAlaThrGluValAlaIleProGln 712  
 Db 2107 CCGCTGAGCTGTACTTGTGTCAAGGTGATGTACGGCGCGGTACGACACCATCCCGCAG 2166  
 Qy 713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
 Db 2167 GACAGGCTACAGAGGTATCGCCACATCAAA---CCCAAGAACAGTGTGTGTGTGTGTGT 2223  
 Qy 733 ArgGlnTyrAlaValAlaArgArgAspSerGlnGlnValHisIleLysSerPheArgArg 752  
 Db 2224 CGTCGATATGCGGTGTCCAGAGAGCCCGCCATGGGACGTCGCCAAGGCGCTTCAGAGC 2283  
 Qy 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisIleGln 772  
 Db 2284 CAGCTCTCACTTACACAGACTCCAGCGCTACATGACAGATTCGTGTGTGTGTGTGTGTGT 2343  
 Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnGlnSerIleSerMetAsn 792  
 Db 2344 GAGACG---AGCCCGCTGAGGAGTCCGTCGTCACTGAGAGACGCTCTCCCTCAAT 2397  
 Qy 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812

Db 2398 GAGCCAGCAGTGGCTCTTCAGACGCTTCTACGCTTCACGTTCAATGTCCACACGCGTGGC 2457  
 Qy 813 IleGlyAspArgCysTyrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 832  
 Db 2458 ATCAAGGGGCAAGTCTTACGTCGATCCAGGAGGAGATCCCGCAGAGGCTTCATCTCCACG 2517  
 Qy 833 LeuLeuCysSerLeuLeuCyAspMetGluAsnLysLeuPheAlaGluValAlaArg 852  
 Db 2518 CTGCTGTGACGCTGT 2577  
 Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisIleuAspGln 872  
 Db 2578 GACGGGTGTCTGT 2637  
 Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
 Db 2638 GCGAAACCTTCTCCAGACACCTCGGTCCAGAGGTGTCCTGAGTATGCTGCGGTGTGAAC 2697  
 Qy 893 LeuGlnLysThrValAlaAsnPheProValGluProGlyThrLeuGlyGluAlaPro 912  
 Db 2698 TTGCGAAGACAGTGTGTGAATTCCTCTGTGAAGACAGAGCCCTGGGTGGACCGGCTTT 2757  
 Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
 Db 2758 GTTCAGATGCGGACCAAGGCTTATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2817  
 Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
 Db 2818 CTGAGAGGTGACAGACACATCACTCACTGACCGGACCTCCATGACAGACAGCTGTACC 2877  
 Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetCysAsnLysLeuLeuSerValLeuArg 972  
 Db 2878 TTCAACCGCGCTTCAAGGTGTGAAGAAATGCGTGTGAACTCTTGGGCTCTTGGCG 2937  
 Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
 Db 2938 CTGAAGGTGACAGACGCTGT 2997  
 Qy 993 AsnIleTyrTyrIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
 Db 2998 AACCTTCAAGATCCCTGT 3057  
 Qy 1013 ProPheAspGlnArgValAlaArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
 Db 3058 CCAATTCATACAGAAAGTTGGAAGAACCCCAATTTTCTCGGCGGTATCTCTGACAGC 3117  
 Qy 1033 AlaSerCysTyrAlaAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
 Db 3118 GCTTCCCTGT 3177  
 Qy 1053 GlySer-----PheProGlnAlaAlaHisIleTrpLeuCysTyrGlnAlaPheLeu 1069  
 Db 3178 GGGCGCGCGGCTCTGCGCTCCGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3237  
 Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 Db 3238 CTCAAGCTGACTCGACACCGGTGTACCTTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3297  
 Qy 1090 GlnLysLeuLeuCysArgLysLeuProGlnAlaThrMetThrIleLeuLysAlaAlaAla 1109  
 Db 3298 CAGACGACGTGAGTGTGAAGGTCTCCGGGAGACACGACTACTGCTGTGAGGCGCGAGCC 3357  
 Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 Db 3358 AACCGGACAGTCCCTCACTCAAGACCATCTGTGAC 3396

RESULT 6  
 BD096291 3396 bp DNA linear PART 27-AUG-2002  
 LOCUS  
 DERIVATION Cells capable of differentiating into myocardial cells.  
 ACCESSION BD096291  
 VERSION BD096291.1 GI:22641879  
 KEYWORDS MO 0148151-A/16.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 3396)  
 AUTHORS Umezawa, A., Hata, J., Fukuda, K., Ogawa, S., Sakurada, K., Gojo, S. and  
 Yamada, Y.  
 TITLE Cells capable of differentiating into myocardial cells  
 JOURNAL Patent: WO 0148151-A 16 05-JUL-2001;  
 KYOWA HAKKO KOGYO CO LTD  
 COMMENT OS Homo sapiens (human)  
 PN WO 0148151-A/16  
 PD 05-JUL-2001  
 PE 27-DEC-2000 WO 2000JP009323  
 PR 28-DEC-1999 JP 99P 372826, 28-FEB-2000 WO PCTJP0001148 PR  
 P1 AKIHIRO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, PI  
 KAZUHIRO SAKURADA, SATOSHI GOJO, YOTI YAMADA  
 PC  
 C12N5/06, C12N5/08, C12P21/08, C12Q1/02, A61K35/28, A61K35/44, A61P9/ PC  
 06', A61P9/04//A61K38/18, C12N5/12  
 CC Cells capable of differentiating into myocardial cells FH  
 KEY Location/Qualifiers  
 FT CDS Location/Qualifiers  
 1. 3396  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 549 a 1157 c 1088 g 602 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,62e-174 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: Gaps: 13

US-09-042-460-2 (1-1122) x BD096291 (1-3396)  
 QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20  
 Db 1 ATCCCGCGCGCTCCCGCTGCGAGCGCGCTGCTGCGCGAGCACTACCGCGAG 60  
 QY 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
 Db 61 GTGCTGCCGCTGCGACGCTTCGTCGCGCGCTGCGCGCGCGCGCGCTGCTGCGAG 120  
 QY 41 ProGlyAspProIysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 Db 121 CGCGGGAGCCCGCGGCTTCGCGCGCTGCTGCGCGAGTGCCTGCTGCTGCGCTGG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuArgGluLeu 80  
 Db 181 GAGCGACGCGCGCGCGCGCGCGCGCGCGCTTCCTCCGACGCTGCTGCTGAAGAGCTG 240  
 QY 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 241 GTGCGCGAGTGTGCGAGAGCTGTGCGACGCGCGCGCGAGAACGCTGCTGCTGCGGC 300  
 QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 Db 301 TTCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGAGGCTTCACACGAGCTGCGC 360  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaTrpMetLeu 140  
 Db 361 AGTACCTGCGCAACAGCTGACCGACGCTGCGCGCGCGCGCGCGCGCTGCTGCTG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysValaLeuTyrLeu 160

421 CTGCGCGCGCTGCGCGAGCGTGTGTTACCTGCTGCGACGCTGCGCGCTTTGTG 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 481 CTGGGCTCTCCACTGCGCTGCGTACAGAGTGTGGGCGCGCGCTGTACAGAGCTCGCGCT 540  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 541 GCCACTGAGCGCGCGCGCGCGCGCGCGCGCGAGCTAGT---GACCGCGAGGCGCTGCGGA--- 594  
 QY 201 AsnPheThrAsnLeuAlaArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 Db 595 -----TGCAGACGCGCGCTGGAACCATGCTGCTGAGAGGCGCGG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 634 GTCCCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTCCAGCGAGT 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 694 CTGCGCGTTCGCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 754 GCGCGAGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTACGAGTACGCTGTTCTGT 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 814 GTGGTGTACCTGCGCAGA-----CCCCCGAGAGACCACTCTTTGGAG 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 859 GGTGGCTCTCTGCGACCGCGCACTCCACCCATCTGCTGCGCGCGCGCGCGCGCGCG 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 919 CCCCACCATCAGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTACGCC 978  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 979 GAGACCAAGCACTTCTTACTTCTCTCAGGACAC---AAGAGACAGCTGCGCGCTCTTC 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyValAlaArgArgLeuValGluIleIlePhe 370  
 Db 1036 CTAAGTCAAGTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTATCTT 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuLysArgThrHisAspLeuSerArgArg 390  
 Db 1096 CTGGGTTCCAGGCGCTGAGTCCAGGAGCTCCCGCAGAGTTGCCCGCGCGCGCGCGCGC 1155  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1156 TACTGGCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTGAGCTGTG 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1216 GCGGTGCTCTCAAGACCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
 QY 431 Leu-----Asn 432  
 Db 1267 GCCGCTGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGGAG 1326  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1327 ACAAGACCCCGTCTGCTGAGTGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTAC 1386  
 QY 453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1387 GGGTGTGCGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCAGG 1446  
 QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
 Db 1447 AACGAAGCGCGCTTCTCTCAGAAACACCAAGATTCACTCTCTGCGGAGAGATGCCAAG 1506

```

QY 493 LeuSerLeuGlnGluLeuMetProlyMetLeuValGluAspCysHisTrpLeuArgSer 512
DB 1507 CTCTGCTGCAGAGACTACGCGGAAGATGAGCGGCGGAGCTGCGCTTGCGGAGG 1566
QY 513 SerProGlyLysAspArgValProAlaIleGlnHisArgLeuArgGluArgIleLeuAla 532
DB 1567 AGCCAGAGGGGTGGCTGGTTCGCGCCGACAGGACCGCTGCTGAGGAGATCTGGCC 1626
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuAspSerPhePheTyr 552
DB 1627 AAGTTCCTGCACTGCTGATGAGTGTGACGTGTCGAGCTGCTCAGGCTCTTTTAT 1686
QY 553 IleThrGluSerThrPheGlnLysAspArgLeuPhePheTyrArgLysSerValTrpSer 572
DB 1687 GTACGGAGACCACTTTCAAAAGAAAGCGCTCTTTTCTACCGAAGAGTGTGAGC 1746
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
DB 1747 AAGTTCGCAAGCATGGAATGACAGACGCTTGAAGAGGGTGCAGCTGCGGAGCTGTCG 1806
QY 593 GlnGluGluValArgHisIleGlnAspThrTrpLeuIleMetProIleCysArgLeuArg 612
DB 1807 GAGCAGAGGTCAAGGACGATCGGAAAGCCAGCCCGCTGCTGAGCTCCAGCTCCG 1866
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
DB 1867 TTCATCCCAAGCTGACGGGCTGGCGCGGATTTGGAACATGACATGACCTGATGGGAGCC 1926
QY 633 ArgAlaLeuGlyArgArgGlyGlnIleGlnHisPheThrGlnArgLeuLysThrPhe 652
DB 1927 AGAAGCTTCGCGAAGAAAGAGGCGCAGCGCTCACCCTGAGGGGTGAAGGCACTGTC 1986
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
DB 1987 AGCGTCTCACTCAGAGCGGCGCGCGCGCGCTCGGGGCGCTGTGCTGGC 2046
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
DB 2047 CTGAGCATATCCACAGGCGCTGGCGCACCTTCGCTGCTGTCGGCGCCAGAGCCG 2106
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
DB 2107 CCGCTGAGCTGATCTTGTCAAGGTGATGTGAGGGCGCGTACAGACATCCCCAG 2166
QY 713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
DB 2167 GACAGGCTCAGGAGGTCATCGCCAGCATCAAA---CCCCAGAACACCTACTGCTG 2223
QY 733 ArgGlnTyrAlaValAlaArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
DB 2224 CGTGGATGCGGTGTCCAGAAAGCCGCCCATGGGCACTGCCAGAGCCTTCAAGAGC 2283
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
DB 2284 CACGCTCTACCTGACAGACTCCAGCGGTACATGACAGAGTTCGCTGCTCAG 2343
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGlnGlnSerIleSerMetAsn 792
DB 2344 GAGACC-----AGCCCGCTGAGGATGCGCTGTGATGAGAGAGTCTCTCCCTGAT 2397
QY 793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
DB 2398 GAGCGCAGCAGTGCTCTTTCAGCGTCTTCAACCTTCATGTGCACACAGCCGTGGC 2457
QY 813 IleGlyAspArgCysTyrThrGlnArgGlnGlyIleProGlnGlySerSerLeuSerThr 832
DB 2458 ATCAGGGGCAAGCCACATCCACATGACGAGGAGATCCCGCAGGGCTCCATCTCTCAGC 2517
QY 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
DB 2518 CTGCTCTGACGCTGTGTCTACGCGACATGAGAAACAGCTCTTTCGGGAGATTCCGCGC 2577

```

---

```

QY 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
DB 2578 GACGGGCTGCTCCCGTGGTGTGGATGATTTCTTGTGTGTGACACTCACCCTCACCCAC 2637
QY 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGlnTyrGlyCysMetIleAsn 892
DB 2638 GCGAAACCTTCTCAGAGACCTGCTCCGAGGTGTCTTGAATAGGCTGCGGTGTGAAC 2697
QY 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyValAlaIlePro 912
DB 2698 TTGGGAGAGCAGAGGTGGAATCTTCCGTGGAAGACAGAGCCCTGGGTGGCAGCGTTT 2757
QY 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
DB 2758 GTTCAGATGCGCGCCACAGCTAATCCCTGTGTGGCGCTGTGCTGTGATACCCGAGCC 2817
QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
DB 2818 CTGAGGTGCAGAGCAGCTACTCTCAGCTATGCCGAGCTCCATCAGACCACTCAGC 2877
QY 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
DB 2878 TTCAACCGCGCTTCAAGGCTGGAGAGAACATGCTGCAAACTCTTGGGGTCTTGCGG 2937
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
DB 2938 CTGAAGTGTCAACCGCTTGTGTGATTCAGGTGAAGCTTCCAGAGGGGTGACCC 2997
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
DB 2998 AACATCTCAAGATCCCTCGCTGACAGGCTGACAGGTTTCAAGCATGTGTGACGCTC 3057
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
DB 3058 CCATTCATCAGCAGAGTTGTGAACACCCACATTTTCTGTGGCGCTCATCTCTGACAGC 3117
QY 1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
DB 3118 GCTTCCCTCTCTACTCTCATCTCAAGGCCAAGACCGAGAGTTCGCTGGGGCCAG 3177
QY 1053 GlySer-----PheProGlnValAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
DB 3178 GGGCGCGCGCGCTGCGCTCCCGAGAGCCGTGACAGTGGCTGTGACCAACAGATTCTG 3237
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
DB 3238 CTCAAGCTGACTCGACACGATGACCTACCTGCGGATCCTCGGGATCAGAGACGCC 3297
QY 3298 CAGACGACGCTGAGTCGGAAGCTCCCGGGAGCAGACGCTGACTGCCCTGGAGCGCGAGCC 3357
DB 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
QY 3358 AACCGGACATGCGCTCAGACTTCAAGACCATCTGTGAGC 3396

```

RESULT 7  
 AX481414 3399 bp DNA linear PAT 16-AUG-2002  
 LOCUS AX481414  
 DEFINITION Sequence 28 from Patent WO02055693.  
 ACCESSION AX481414  
 VERSION AX481414.1 GI:22316328  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.  
 Method for inhibiting the expression of a target gene  
 Patent: WO 02055693-A 28 18-JUL-2002;  
 Ribopharma AG (DE)  
 FEATURES  
 Location/Qualifiers

```

source      1..3399
organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
SE COUNT    550 a 1157 c 1089 g 603 t
IGIN

ignment Scores:
ad. No.:      3,628-174      Length:      3399
pre:          3505.00      Matches:      719
cent Similarity: 72.94%      Conservative: 122
st Local Similarity: 62.36%      Mismatches: 260
ary Match:      59.40%      Indels:      52
:               6          Gaps:      13

-09-042-460-2 (1-1122) x AX481414 (1-3399)
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
1 ATGCGCGGCGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 60
21 ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValcIn 40
61 GTGCTCCGCTGCGCCAGTTCGTGCGGCGCTGCGGCGCCCGAGGCTGCGCGCTGGTGG 120
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
121 CGCGGGAGACCGCGCGCTTCGCGCGCTGCTGGCCAGTGCCTGCTGGTGGTGGCTGG 180
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyGluLeu 80
181 GAGCGACGGCGCGCGCGCGCTCTTCCGCGAGGTGCTGCTGAGGAGGTG 240
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
241 GTGCGCGGAGTGTGCGAGAGCTGTGCGAGCGCGTGGCGGAGAGCTGCTGGCTTCGG 300
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
301 TTCGCGCTGTGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 360
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
361 AGTACCTGCGCCACACGCGTACCGACGACCTGCGGCGGAGCGGCGCGCTGGCGCTG 420
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
421 CTGCGCGCGCTGCGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
481 CTGCTGCTCCAGCTGCGCTACCAAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCT 540
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
541 GCACACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
595 -----TGCAGACGGCGCTGGAACCATAGCGCTCAGGAGGCGCGG 633
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrThrSer 240
634 GTCCCTCTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 693
241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGlyPro--- 259
694 CTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 753
260 -----HisArgGlnValLeuProThrProSerClyLysSerTyr--- 272
754 GGGCAGGGGTCTCTGGGCGCCACCGCGGAGGACGCTGGAGCCGAGTACCGTGGTTCGT 813

```

```

QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
DB 814 GTGTGTACCTGCCAGA-----CCGCGGAGAGACCCACCTCTTTGGAG 858
QY 293 GlyLysValSerAspLeuSerLeuSer--GlySerValCysCysLysHisLysProSer 311
DB 859 GGTGGCTCTCTGGCAGCGCCACTCCACCCATCGGTGGCGCGCCAGCACCGCGGCG 918
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 919 CCCCCATCCACATCGCGGCCACACGCTCCCTGGGAGACGCGCTTGTCCCGGTGTACGCC 978
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGlnLeuArgLeuAsnProSerPhe 350
DB 979 GAGACCAAGCACTTCTCTACTCTCTCAGGCGAC---AAGGAGCAGCTGCGGCGCTCTTC 1035
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyValArgArgLeuValGluIleIlePhe 370
DB 1036 CTACTCAGCTCTCTGAGCGCCGAGCGCTGACTGGCGCTGGAGGCTCTGGAGACCATCTTT 1095
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
DB 1096 CTGGGTTCCAGCGCCCTGGATGCCAGGACTCCCCGCGAGTTGCCCGCTGCCCGACGCG 1155
QY 391 TyrTyrGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
DB 1156 TACTTGGCAATGCGCGCCCTGTTTCTGGAGCTGCTGGGAACACACGCGCAGTGCCTTAC 1215
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
DB 1216 GGGGTGCTCTCAAGACGCACTGCCGCTGGAGCTGCG-----GTACCCCGACGA 1266
QY 431 Leu-----Asn 432
DB 1267 GCGGTGTCTGTGCCGCGGAGAACCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAG 1326
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452
DB 1327 ACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
QY 453 GlyPheLeuAlaGlyCysLeuValSerAlaSerLeuTyrGlyThrArgHis 472
DB 1387 GACTTCTGTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
QY 473 AsnGluArgArgPhePheLysAsnLeuLysPheIleSerLeuGlyLysTyrGlyLys 492
DB 1447 AACGACCGCGCTTCTCAGGACACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAG 1506
QY 493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512
DB 1507 CTCTGCTGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlyLeuAla 532
DB 1567 AGCCAGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
QY 533 ThrPheLeuPheTyrLeuMetAspThrTyrValGlnLeuLeuArgSerPhePheTyr 552
DB 1627 AAGTCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572
DB 1687 GTACGAGAGACCACTTCAAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1746
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
DB 1747 AAGTTGCAAGCATTTGGAATCAGACAGCATGAGAGGGTGCAGCTGCGGAGCTGTGCG 1806
QY 593 GlnGluGluValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgLeuArg 612
DB 1807 GAAGCAGAGGTTCAGGACGATCGGAGCGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1866
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632

```

```

1867 TTTATCCCCAGCCTGACGGGCTGGGCGCATTTGTGAACATGAGCTACCTGCTGGGAGCC 1926
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1927 AGAAGTTTCGACAGAGAAAGAGCGCGCGCTCACCTCGAGGGTGAAGCACTGTTC 1986
653 SerMetLeuSerThrGluArgThrLysHisPheProHisLeuMetGlySerSerValLeuGly 672
1987 AGCGTGTCAACTACGACGCGCGCGCGCGCGCTCTCTGGCGCGCTCTGTGTGGGC 2046
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2047 CTGACAGATATCCACAGGGCTGGCGACCTTGTGCTGGTGGTGGCGCCAGACCCG 2106
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2107 CCGCTCGAGCTGTACTTTGTCAAGGTGGATGTACGCGGCGCGGTACACACCATCCGCCAG 2166
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2167 GACAGGCTCAGGAGGTATCGCCAGCATCATCAAA---CCCGAAGACAGTACTGGTG 2223
733 ArgGlnTyrAlaValAlaArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2224 CGTCGGTATGCGGTGTCAGAGGCGCGCCCATGGGCACGCTCGCAAGGCTTCAAGAGC 2283
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2284 CACGTCTCTACCTTGACAGACCTCCAGCCGTATCATCGACAGTTCGTGGCTCACCTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2344 GAGACC-----AGCCCGCTGAGGATGCGGTGTCATCGAGCAGAGCTCTCCCTGAT 2397
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812
2398 GAGCCACGAGTGGCTCTTCGAGCTCTCTACGCTTCATGTGCCACACCGCGTGGC 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2458 ATCAGGGGAGTCTTACGTCCAGTGCAGGGGATCCCGAGGAGCTCCATCTCTCCAG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2518 CTGCTCTGACGCTGTCTACGGCGCATGAGAACTGAGTGTTCGGGGATTCGGCG 2577
853 AspGlyLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872
2578 GACGGGCTGCTCCTCGCTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2638 CGCAAAACCTTCTCCAGACCTGTCCGAGGTGTCCTGAGTATGCTCGTGGTGAAC 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2698 TTGGGAAGACAGTGGTGAATCTCCCTGTAGAAGACAGGCGCTGGGTGGCACGGCTTT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2758 GTTCAGATGCGGCGCCACGGCTATTCTCCCTGGTGGCGCTGTGTGTGATACCCGACC 2817
933 LeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2818 CTGGAGGTGCAGAGCGACTATCCAGGTATGCCCGGACCTCCATCAGAGCCAGTCTCAC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2878 TTCAACCGCGCTTCAAGCTGGAGGAAATGCTGCGCAAACTCTTTGGGTCTTGGCG 2937
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992

```

```

Db 2938 CTGAGTGTACAGCCTGTTTCTGATTGCGAGTGAACAGCTCCAGCGTGTGCACC 2997
Qy 993 AsnIleTyrIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
Db 2998 AACATCTACAGATCCTCCTGCTGCGAGCGTACAGGTTTACGATGTGTGTGAGCTC 3057
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3058 CCATTTCATCAGCAAGTTTGAAGAACCCCAATTTTCTGCGCGTCACTCTTGACAGC 3117
Qy 1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaser 1052
Db 3118 GCCTCCTCTGCTACTCATCTCTGAAGACCAAGACGAGGATGTCTGCGGGCCCAAG 3177
Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
Db 3178 GCGCGCGCGCGGCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3237
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
Db 3238 CTCAGCTGACTCGACACCGTGTCACTACGTGCCACTCTGGGTCACTCAGGACAGCC 3297
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
Db 3298 CAGACGAGCTGAGTCGGAAGCTCCGCGGAGCAGCGTGACTGCTGCGCGCGCGAGCC 3357
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3358 AACCCGGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3396

RESULT 8
ARI04587
LOCUS ARI04587 4015 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 224 from patent US 6093809.
ACCESSION ARI04587
VERSION ARI04587.1 GI:12817295
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4015)
AUTHORS Cech,T.R. and Lingner,J.
TITLE Telomerase
JOURNAL Patent: US 6093809-A 224 25-JUL-2000;
FEATURES
1. .4015
source /organism="unknown"
BASE COUNT 663 a 1363 c 1275 g 714 t
ORIGIN

Alignment Scores:
Pred. No.: 4,54e-174 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
Query Match: 59.40% Indels: 52
DB: 6 Gaps: 13

US-09-042-460-2 (1-1122) x ARI04587 (1-4015)
Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115
Qy 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGCTGGCCACGTTCTGTGGCGGCTTGGCGGCGCCAGGCTGGCGCTGGCGAG 175
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
Db 176 CGCGGAGACCCCGCGGCTTTTCGCGCGCTGTGGGCGGCTGCTGCTGCTGCTGCTGCTG 235
Qy 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

```



1271 GGGGTGCTCTCAAGACGACTGCCCGCTGCAGAGTGG-----GTCACCCCGACA 1332

431 Leu-----Asn 432  
:::

1322 GCCGGTGCTGTGCGCGAGAAAGCCCAAGGGCTCTGTGGCGGCCCGCAGAGGAGTC 1381

433 ThrSerProProlHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTy 452  
|||:::  
1382 ACAGACCCTCGTGGCTGGTGACGTCTCCGCCACACAGCAGGCCCTTGGAGGTTGAC 1441  
|||:::  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
|||:::  
1442 GCCTTCGTGGCGGCTGCTGCTGGCGGCTGTGCTCCCACAGGCTCTGGGGCTCCAGGCAC 1501  
|||:::  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyRcLys 492  
|||:::  
1502 AACGAACCCCGCTTCCTCAGACACACCAAGAATTCTATCTCCCTGGGAAGCATGCCAAG 1561  
|||:::  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
|||:::  
1562 CTTCTGCTGCAGGAGCTGACCTGGAGATGAGCGTGCAGGACTGCGCTTGGCTGCACAG 1621  
|||:::  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
|||:::  
1622 AGCCAGCGGGTTGGCTGTGTTTCGGCGCAGACACCGTCTCGCTGAGAGATCTCGGCC 1681  
|||:::  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTy 552  
|||:::  
1682 AGTTCTCTGCACCTGGCTGATGAGTGTGTACTGCTGCAGCTGCTCAGGTCTTCTTTTAT 1741  
|||:::  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyArgLysSerValTrpSer 572  
|||:::  
1742 GTCCAGGAGACCACTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
|||:::  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
|||:::  
1802 AAGTTGCAAGATTTGGAATCAGACACACTTGAGAGGGTGCAGCTCGGGAGGTGTGCG 1861  
|||:::  
593 GlnGluValValArgHisHisGlnAspThrTrpLeuAlaMetProileCysArgLeuArg 612  
|||:::  
1862 GAAGCAGAGGTCAAGCAGCATCGGAGACCCAGCGCCCTGCTGAGCTCCAGACTCCGC 1921  
|||:::  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTySerMetClyThr 632  
|||:::  
1922 TTCTATCCCCAAGCTTCAGGGCTCGCGGCCGATTTGACATGCATAGCTGTGTGGAGGCC 1981  
|||:::  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
|||:::  
1982 AGNAGCTCCGACAGAAAAGAGGCGCGAGCGGTCTCACCTCGAGGGGTGAAGCACTGTC 2041  
|||:::  
653 SerMetLeuAsnTyRcGluArgThyLysHisPheHisLeuMetClySerSerValLeuGly 672  
|||:::  
2042 AGCGTCTCAACTAGCAGCGCGCGCGCGCGCGCGCTCTCTGTGGCGCCCTTGTGTGTGGC 2101  
|||:::  
673 MetAsnAspIleTyArgThyTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
|||:::  
2102 TGGACGATATCCAAGGCGCTGCGGCACCTTGTGTGTGGTGTGGCGGCCCGAGACCG 2161  
|||:::  
693 ThrProArgMetTyPheValLysAlaAspValThrGlyValatyRaspAlatleProGln 712  
|||:::  
2162 CGCCCTGACCTGACTTTGTCAAGTGSATGTGACGGCGCGGTACGACACCATCCCCCAG 2221  
|||:::  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyRcValle 732  
|||:::  
2222 GACAGGCTCACGAGGTTCATCGCAGCATCATCAAA---CCCAGAACACGACTGTGCGTG 2278  
|||:::  
733 ArgGlnTyRcAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArg 752  
|||:::  
2279 GTGCGGTATGCGTGTGTCAGAGGCCGCCCATGGGCACTCGCAGAGCCCTTCAAGAGC 2338  
|||:::  
753 GlnValThrThrLeuSerAspLeuGlnProTyRcMetGlyGlnPheLeuLysHisLeuGln 772  
|||:::  
2339 CAGCTCTTACTTTCAGAGACTCGACCGGTACATGCGACAGTTGCGCTCTCTGTGAG 2398  
|||:::

```
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerMetAsn 792
      :::::
2399 GAGACC-----AGCCCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCCCTCGAAT 2452
      :::::
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812
      :::::
2453 GAGGCCAGCAGTGGCCCTCTTCAGCGTCTTCTACGCTTCATGTGCCACCGCGGTGGC 2512
      :::::
813 IleGlyAspArgCysTyThrGlnCysGlnGlyLeProGlnGlySerSerLeuSerThr 832
      :::::
2513 ATCAGGGCAAGTCTCATGCTCAGTCAGGAGATCCCGCAGGAGCTCCATCTCTCCAGC 2572
      :::::
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
      :::::
2573 CTGCTCTGACGCTGTGCTCGGCGCATGAGAACAGCTGTTCGGGGATTCGGCGG 2632
      :::::
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
      :::::
2633 GACGGGCTGCTCTGCGTTCGTTGGTGGATGATTTCTTTGTTGGTGCACCTCACCCTCACCCAC 2692
      :::::
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrglyCysMetIleAsn 892
      :::::
2693 GCGAAACCTTCTCAGACCCCTGGTCCGAGGTGTCCTCAGTATGCTGCTGATACCCGAGC 2752
      :::::
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
      :::::
2753 TTGCGGAAGACAGTGGTGAATCTTCTCTGTAGAACGAGGCGCTGGGTGGCAGCGCTTTT 2812
      :::::
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
      :::::
2813 GTTCAGATGCGCGGCCACCGGCTATTCCCTGGTGGCGCTGCTGCTGATACCCGAGC 2872
      :::::
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
      :::::
2873 CTGAGGTGCGAGCGACTACTCAGCTATGCGCGACCTTCATCAGACCGACTCTACCC 2932
      :::::
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
      :::::
2933 TTCAACCGCGCTTCAAGGCTGGGAGGAACATGCTGCGCAAACTCTTTGGGTCTTGGCG 2992
      :::::
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
      :::::
2993 CTGAAGTGTACAGCGTCTTCTCGATTTGAGGTGAACGCTCCAGACGGTGTGCACC 3052
      :::::
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
      :::::
3053 AACATCTACAAGATCTCTCTGCTCAGCGGTACAGTTTTCAGCATGTGTGTCTGACGCTC 3112
      :::::
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyLysIleSerSerGln 1032
      :::::
3113 CCATTTTCATCAGCAAGTTTGAAGAACCACCATTTTCTCGCGCTCATCTCTGACAGC 3172
      :::::
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
      :::::
3173 GCCTCCCTCTCTACTCTCATCTCTGAAAGCCAGAACCCAGGAGTGTGCTGGGGGCCAAG 3232
      :::::
1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
      :::::
3233 GCGCGCGCGCGCCCTCTGCGCTCCGAGCGCGCTGAGTGGTGTGCCACCAAGCATTTCCCTG 3292
      :::::
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
      :::::
3293 CTCAGCTGACTCAGACCGGTGTCATCTAGTCCCATCTCTGGGGTTCATCTCAGGACAGCC 3352
      :::::
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
      :::::
3353 CAGACCGAGCTGAGTCGAAAGCTCCCGGGGACGAGCTGACTCTCCCTGAGGCGCGAGCC 3412
      :::::
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
      :::::
3413 AACCCGGCAGTGGCTCTCAGACTTCAAGACCATCTCGAC 3451
```

```
RESULT 9
ARI175848
LOCUS ARI175848 4015 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 224 from patent US 6309867.
ACCESSION ARI175848
VERSION ARI175848.1 GI:17917147
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4015)
AUTHORS Cech,T.R. and Nakamura,T.
TITLE Telomerase
JOURNAL Patent: US 6309867-A 224 30-OCT-2001;
FEATURES Location/Qualifiers
            1..4015
                /organism="unknown"
BASE COUNT 663 a 1363 c 1275 g 714 t
ORIGIN
Alignment Scores:
Pred. No.: 4,54e-174 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
Query Match: 59.40% Indels: 52
DB: 6 Gaps: 13
US-09-042-460-2 (1-1122) x ARI175848 (1-4015)
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGGCGCTCCCGCTGCCGAGCGGTGCTGCTGCTGCGCAGCCTACCGCAG 115
QY 21 ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGGTGCGCACGTTGCTGCGCGCGCTGCGGCGCCCGCGGCTGCGGTGTCAG 175
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
Db 176 CGCGGGGACCGCGGCTTTCCCGCGCGTGTGTGGGCCAGTGTGCTGCTGCGTCCCTGG 235
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGlnLeu 80
Db 236 GACGACGCGCGCGCGCGCGCGCTCTCTCCCGCAGGTGCTCTGCTGCTGAAGAGCTG 295
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
Db 296 GTGGCCCGAGTGTGTGAGAGGCTGTGCGAGCGCGCGGCGGAGAACGTGCTGCGCTTCG 355
QY 101 PheGluLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
Db 356 TTGCGGTGTGTGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
Db 416 AGCTACTGCTGCCAACCGGTGACCGCAGCAGTGTGCGGGGAGCGCGGCGGTGGGGGTG 475
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
Db 476 CTGCGCGCGTGGCGAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
Db 536 CTGGTGGCTCCCAAGTGTGCGCTACCGAGGTGTGCGGCGCGCGCGCGCGCGCGCGCG 595
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
Db 596 GCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerSerArgGlnGlnAlaPro 220
Db 650 -----TGCGAACCGGCGCTGGAAACCATAGCTGAGGAGCGCGCGG 688
```

221 LysProLeuAlaLeuProSerArgGlyThrIysArgHisLeuSerLeuThrSerThrSer 240  
241 ValProSerAlaIysLysAlaArgCysTrpProValProArgValGluGluGlyPro--- 259  
749 CTGGCGTTGCCAAAGAGCCAGAGCGTGGCGCTCCCTCAGCGGAGCGAGCCCGCTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGTCTCGGGCCCAACCGGGCAGAGCGGTGGACCGAGTACCGTGTTCGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGTGTCACTGCCAG-----CCCGCCGAAGAAGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGGCTCTCGGCACGGCCACTCCACCCATCCGTGGGCCGCCAGCACCGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAenAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCATCCGCGCCACACGTCCTCGGACACGCTTGTCCCGCGGTATCGCC 1033  
331 GluThrArgHisPheLeuTySerArgGlyAspGlyGlnGluArgLeuAenProSerPhe 350  
1034 GAGACAGACCTCTCTACTCTCTCAGCGAC---AAGAGCAGCTCGCGCTCTTC 1090  
351 LeuLeuSerAsnGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCAGGCCACCTGCTGCGCTCGAGGCTCGTGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTCCAGGCCCTGGATGCCAGGACTCCCCGAGTTCGCCCGCTGCCCGCCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAenHisAlaGluCysGlnTyr 410  
1211 TACTGCCAATCGGCCCTCTTCTGGAGCTGTCTGGGAACCCAGCAGCGAGTGCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAenGlnValThrAspAla 430  
1271 GGGGTCTCTCAAGACGCACTGCCCGCTGGAGTGG-----GTACCCACGCA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTCCCGGAGAACCCAGGCTCTGTGGCGGCCCGGAGAGGAC 1381  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCTGGTGCAGCTGTCTCCGCCAGCACAGCAGCCCTGGCAGGTATC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGTGGCGGCTCTCTGGCGCGCTGGTGGTGGCCCGAGGCTCTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyGlyLys 492  
1502 AACGACGCGCTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGAGAGCTGAGTGGAAAGATGAGCGTGGGACTGCGCTTGGCTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGTGGCTGTGTTCCGCGCGCAGAGCACCGTCTCGTGAGAGATCCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGTGTAGTGTACGTGCTGAGTGTCTCAGGTCTCTCTTTAT 1741

553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGAGACCAAGTTCAAAAGAACAGCGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGACATTGGAATCAGACAGCACTTTGAAGAGGTGCAGCTCGGAGGTGTCG 1861  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGTTCAGGAGCATCGGGAAGCAGCGCCCTCTGCTACGTCCAGTCCAGTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCTGACGGCTCGCGCGATTTGAAACATGACTACGTCTCGTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGCGAGAGAAAGAGGCGCGAGCTCTCACTCGAGGGTGAAGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGCGCGCGCTCTCTGGGCGCTCTCTGTGTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCTCGCGCACCTTCTGTCTGTGTGGGCGCCAGGACCGC 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGluValTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTTGTCAAGTGTGATGACGGGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGCTTCAGGAGGTATCGCCAGCATCATCAA---CCCCAGAACACGACTCGTG 2278  
733 ArgGlnTyrAlaValAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGATGCGGTGGTCCAGAAAGCGCCCATCGGCACGTCGCGAGGCTTCACAGAC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACTTTCAGACACCTCCAGCGCTACATGCAGTTCGTGGTCACTTCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGAC-----AGCCCGCTGAGGATCGCTCATCGACGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGACATGCGCTTCGACGTCTTCTACGCTTCATGTGCCACACCGCGGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTCAGTCCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPhePheLysAspMetGluAsnLysLeuPheAlaGluValArg 852  
2573 CTGCTCTGACGCTGTCTACGGCGATGAGAACAGCTGTTTGGCGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuLeuAspGln 872  
2633 GACGGCTGCTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCCAGACCTGTCCGAGGTCTCTGAGTATGGCTCGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaPro 912  
2753 TTGCGGAGACAGTGTGAATCTTCCCTGAGAACAGACGAGGCGCTGGTGGTGGTGGT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932

```

2813 GTTCAGATCGCGCCACCGCTATTCCCTGCTGGCGCTGCTGCTGATACCGGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGAGAGTGCAGAGCAGTACTCCAGCTATGCGCGGAGCTCCATCAGAGCCAGTCTCACC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972
2933 TTCAACCGGGCTTCAAGCTGGGAGAACATGCTCGCAACTCTTTGGGGTCTTCGGG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGACCTGTTCTGATTTGCAAGTGAACAGCTCCAGACGCTGTGCACC 3052
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAGATCCTCTGCTGAGCGGTACAGTTTCAGCATGTGCTCGAGCTC 3112
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTCGCGCTCATCTCTGACACG 3172
1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCTCCTCTGCTACTCTCATCTGAAAGCCAAAGACGAGGAGTGTGCTGGGGGCCAAG 3232
1053 GlySer-----PheProGlnAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGCGCGCGCGCCCTGCGCTCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAGCTGACTCGACACCGCTGCACCTACGTGCGACCTCTCGGGGTCACTCAGGACAGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGlnAlaThrMetThrIleLeuLysAlaAla 1109
3353 CAGACCGACTGAGTCGGAAGCTCCCGGGAGCAGCGCTGACTGCCCTGGAGGCGCAGCC 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCCGGCTGCGCTTCAGACTTCAAGACCATCTCTGGAC 3451

```

```

SPLIT 10
182221
CUS
FINITION
Sequence 1 from patent US 6337200.
CESSION
AR182221
REASON
AR182221.1 GI:20225137
WORDS
Unknown.
URCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 4015)
AUTHORS
Morin, G.B.
TITLE
Human telomerase catalytic subunit variants
JOURNAL
Patent: US 6337200-A 1 08-JAN-2002;
AURES
Location/Qualifiers
source
1. .4015
/organism="unknown"
SE COUNT 663 a 1363 c 1275 g 714 t
IGIN
Ignment Scores:
ed. No.: 4,54e-174 Length: 4015
ore: 3505.00 Matches: 719
cent Similarity: 72.94% Conservative: 122
st Local Similarity: 62.36% Mismatches: 260
ery Match: 59.40% Indels: 52
: 6 Gaps: 13
-09-042-460-2 (1-1122) x AR182221 (1-4015)

```

```

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGGTGCGCTCCCTGCTGGCAGCCACTACCGCGAG 115
QY 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGCTGGCCACGTTCTGCGCGCGCTGGCGCCCGAGGCTGGCGCTGGTGCAG 175
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
Db 176 CCGCGGGACCGCGCGGCTTTCCGCGCGCTGCTGGCCAGTGCCTGGTGGCTGCCCTGG 235
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
Db 236 GACGACCGCGCGCGCGCGCTCTCTTCGCCAGGTGCTCTGCTGAAGAGAGCTG 295
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
Db 296 GTGGCGCGAGTGTGCAAGGCTGTGCGAGCGCGCGCGAGAACGTCGTGGCTTCGGC 355
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
Db 356 TTCGCGCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACCCAGCGTGGCG 415
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaThrMetLeuLeu 140
Db 416 AGCTACCTGCGCAACAGGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGCGTCTG 475
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
Db 476 CTGCGCGCGGTGGCGGAGAGTGTGCTGCTGCTGCGAGCGCTGCGCGCTCTTGTG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
Db 536 CTGTGTGCTCCCGCTGCGCTACAGGTGTGGGGCGCGCGCTGTACAGCTCGCGCT 595
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
Db 596 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
Db 650 -----TGGAAAGCGCGCTGGNACCATAGCTCAGGAGCGCGG 688
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
Db 689 GTCCCGCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
Db 749 CTGCGCTTGCACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
Db 809 GGGAGGCGGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
Db 869 GTGGGTGCACCTGCCAGA-----CCCCCGAAGAACCGACCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311
Db 914 GGTGGCTCTCTGGCAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro-----PheIle 330
Db 974 CCCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370

```

1091 TCTACTCAGCTCTCTGAGGCGCCAGCGCTGCTGGCGCTCGAGAGCTCGTGGAGCACCATTTT 1150

371 LeuGlySerArgProArgThrSerGlyProLeuCyArgThrHisArgLeuSerArgArg 390

1151 CTGGGTTCCAGGCGCTCGATGTCACAGGACTCCCGCGAGTTCGCCCGCTGCCCCAGGCG 1210

391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCyArgTyr 410

1211 TACTGGCAAAATGCGCGCCCTCTGTTCTCGGAGCTGCTTGGGAACACAGCGCAGCTGCCCTAC 1270

411 ValArgLeuLeuArgSerHisCyArgPheArgThrAlaAsnGlnGlnValThrAspAla 430

1271 GGGGTGCTCTCCAGAGCGCACTGCCCGCTCGAGCTGCG-----GTCCGCCAGCA 1321

431 Leu-----Asn 432

1322 GCCGCTGTCTGTCGCGGAGAGAGCCCGCAGGGCTCTGTGGCGGCCCGCAGGAGGAGAC 1381

433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452

1382 ACAGACCCCCCTGCCCTGGTGGAGCTGCTCCGCGACACAGCAGCGCCCTGCGCAGGTGAC 1441

453 GlyPheLeuArgAlaCyLeuCyLysValValSerAlaSerLeuTrpGlyThrArgHis 472

1442 GGTCTGTGGGCGCTGCCCTGGCGGCTGGTGGCCCGAGGCTCTGGGGTCCAGGCAC 1501

473 AsnGluArgArgPhePheLysAsnLeuLysPheLysLeuGlyLysTyrGlyLys 492

1502 AACGAACCGCGCTCTCTCAGGAAACCAACAGAGTTCTATCTCCCTGGGGAACATGCCAAG 1561

493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCyHisTrpLeuArgSer 512

1562 CTCTCGCTGCGAGAGCTGACGTGGAGATGAGCGTGGGAGATGCGCTTGGCTGGCGGAGG 1621

513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532

1622 AGCCAGGGGTGGCTGTGTCTCGGCGCGCAGAGACCGCTCGCGTGGAGAGATCTCGGCC 1681

533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552

1682 AAGTTCCTGCACTGCTGATGAGTGATCGTCTCGAGCTGCTCAGGTCTTCTTTTAT 1741

553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572

1742 GTCAGGAGACCGTTCAAAAGAACACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801

573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592

1802 AAGTTGCAAGCATTTGGAATCAGACACATTCAGAGAGGTGCAGCTCGGAGAGCTGCG 1861

593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCyArgLeuArg 612

1862 GAACAGAGGTCAGGACGATCGGGAAGCCAGCGCCCTGCTGACGCTCCAGACTCCGC 1921

613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632

1922 TTCACTCCCAAGCTGACGGCTCGCGCGCATTTGGAACATGCATCTAGCTGTGGAGACC 1981

633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652

1982 AGAAGTCCCGCAGAAAAGAGGGCCGCGGCTCTCACTCGAGGTGAAGCACTGTTC 2041

653 SerMetLeuAsnTyrGluArgThrLysHisProHisIleLeuMetGlySerSerValLeuGly 672

2042 AGCGTCTCAACTACAGCGCGCGCGCGCCCGGCTCTCTGGGGCGCTCTGTGTGGCG 2101

673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692

2102 CTGGAGCATATCCACAGGCGCTGCGCACTTGTGCTGGTGTGGGCGCCAGGACCGC 2161

693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValArgAspAlaIleProGln 712

2162	CGCGCTGAGCTGACTTTGTCAAGGTGGATGTACGGGGCGGTAGCACACCATCCCCAG	2222
713	GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle	732
2222	GACAGGCTCAGGAGGTCTATCGGCAGCATCATCAAA---CCCGAGAACACGTACTGGGTG	2278
733	ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg	752
2279	GTCGGTATGCGGTGTCAGAAAGCCGCCATGGCGACGTCGCAAGGCCCTTCAAGAC	2338
753	GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln	772
2339	CACGTCCTCTACCTTGACAGACCTCCAGCCGTCATATGCGACAGTTCGTGGCTCACCTGCAG	2398
773	AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn	792
2399	GAGACC-----AGCCCGCTGAGGGATGCGCTCGTCATCGACAGAGCTCTCTCCCTGAAT	2452
793	GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys	812
2453	GAGGCCAGCAGTGCCTCTTCGACGCTCTTCCTACGCTTCATGTGCGCACACGCCGTGCGC	2512
813	IleGlyAspArgCysTyrThrGlnCysGlnGlyLeProGlnGlySerSerLeuSerThr	832
2513	ATCAGGGCAAGTCTCTACGTCCAGTCCAGAGGGATCCCGCAGGGCTCCATCCTCTCCACG	2572
833	LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg	852
2573	CTGCTCTCGACCCCTGCTACGGCGACATGGAGACACAGCTGTTTCGGGGATTCGGCGG	2632
853	AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln	872
2633	GACGGCTGCTCCTCGTTCGTTGGTGGATGATCTTCTTGTGGACACCTCACCTCACCCAC	2692
873	AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn	892
2693	GGCAAAACCTCTCTCAGACCCCTGGTCCGAGGTGTCCTTAGATGGCTGCGTGGTGAAC	2752
893	LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro	912
2753	TTCGGAGAGACAGTGGTGAACCTCCCTGTAGAAGACGAGGCCCTCGGTGGCACGGCTTTT	2812
913	TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr	932
2813	GTTTCAGATCGCGGCCACAGCCCTATTCCCTCGTGGCGCTGCTCTCGATACCCGGACC	2872
933	LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr	952
2873	CTGGAGGTGCAGCGGCTACTTCAGCTATGCGCGACCTCCATCAGAGCCAGTCTCAC	2932
953	PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg	972
2933	TTCACACCGCGCTTCAGAGGTGGAGAGAACATGCGTCGCAAACTCTTTGGGGTCTTCGG	2992
973	LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle	992
2993	CTGAAGTGTACACACCTGTTCTTGATTTTCAGGTGACAGAGCTCTCAGACGGTGTGACC	3052
993	AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu	1012
3053	AACATCTACAAGATCCTCTGCTGTCAGGCGTACAGGTTTCACGCATGTGTGTCAGCTC	3112
1013	ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln	1032
3113	CCATTTTCATCAGCAAGTTTGGAAGACCCACATTTTCTCTCGCGCTCATCTCTGACACG	3172
1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyWetThrLeuLysAlaSer	1052
3173	GCTCTCCTCTCTACTCCATCTCTGAAGCCAGAACGACGGATGTCTGCTGGGGCCAA	3232
1053	GlySer-----PheProProGluAlaAlaHisTyrPleuCysTyrGlnAlaPheLeu	1069
3233	GGCGCGCGCCCTCTGCTCCCTCCGAGCGCGTACGTGCTGTGCGCACAGACATTCCTG	3292

1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 3293 CTCAGCTGACTCGACACCGTGTACCTACGTGCGACCTCTCTGGGGTCACTCAGGACAGCC 3352  
 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
 3353 CAGACCGCACTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGCGCGAGCC 3412  
 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 3413 AACCCGGCACTGCCCTCAGACTTCAGACCACTCTCTGGAC 3451

## SULF 11

224455  
 CUS AR224455 4015 bp DNA linear PAT 26-SEP-2002  
 FINITION Sequence 1 from patent US 6440735.  
 CESSION AR224455  
 RSION AR224455.1 GI:23333293

WORDS  
 ORCE  
 ORGANISM

PERENCE  
 AUTHORS

TITLE

JOURNAL  
 ATURES

1. .4015  
 /organism="unknown"

SE COUNT 663 a 1363 c 1275 g 714 t

IGIN

## ignment Scores:

ed. No.: 4,54e-174 Length: 4015  
 ore: 3505.00 Matches: 719  
 rent Similarity: 72.94% Conservative: 122  
 st Local Similarity: 62.36% Mismatches: 260  
 ery Match: 59.40% Indels: 52  
 : 6 Gaps: 13

-09-042-460-2 (1-1122) x AR224455 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATCCCGCGGCTCCCGCTGCCAGCGGTGCGCTCCCTGCTGGCGAGCACTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
 116 GTGCTCGCTGGCCACGTTCTGCGGCGCTGGCGGCGCCAGGGCTGGCGGCTGGTGCGAG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGCGGGGACCGCGGCTTTCGCGGCTGGTGGCGGAGTGCCTGGTGGCTGCCCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GACGACGCGCGCGCGCGCGCGCTCTCCCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTG 295  
 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGGCCGAGTGTGTCAGAGGTGTGTCGAGCGCGCGCGCGAGCAAGCAAGTGTGCTGCGC 355  
 101 PheGluLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGCTGGACGGGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGCTACTGTCACCAACGCTGACCGACGACGCTGCGGGGAGCGGGGCGCTGGGGGCTGCTG 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160

Db 476 CTGCGCGCGGTGGCGCGACGAGCTGCTGTGTTTTCACCTGCGACGCTGCGCGCTTTTGTG 535  
 Qy 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 536 CTGGTGGCTCCAGCTGGGCTTACAGGTGTGGGGCGCGCTGTACAGCTCGCGCT 595  
 Qy 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 596 GCCACTAGCGCGCGCGCGCGCGCACTAGT---GGACCCGGAAGCGCTCTGGGA--- 649  
 Qy 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
 Db 650 -----TGGAACGGCGCTGGAAACCATAGCTCAGGAGGCGCGG 698  
 Qy 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 689 GTCCCGCTGGGCTTGCAGCGCGCGGTGCGAGGCGCGGGGCGAGTGCAGCCGAAAT 748  
 Qy 241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGlyPro--- 259  
 Db 749 CTGCGTTCGCAAGAGCGCGCGGCTGGCGCTGCGCTGAGCGCGAGCGGCGCGCTT 808  
 Qy 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 809 GGGCAGGGGTCTCTGGCGCCACCGCGCGAGCGCGTGCACCGAGTGCAGCTGTTCTGT 868  
 Qy 273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerSerLys 292  
 Db 869 GTGGTGTCTCCTGCCAGA-----CCCGCCGAAAGACCACTCTTTGGAG 913  
 Qy 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 914 GGTGGCTCTCTGGCAGCGCGCACTCCACCCATCTCGTGGCGCGCGAGCACCGCGGGC 973  
 Qy 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 974 CCCCCATCCACATCGCGCGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTAGGCC 1033  
 Qy 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 1034 GAGACCAAGCACTTCTCTCTCTCTCAGGGCAGC---AAGGAGCAGCTGGCGGCTCTCTTC 1090  
 Qy 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 Db 1091 CTACTCAGCTCTCTGAGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150  
 Qy 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 Db 1151 CTGGGTTCAGCGCGCTGGATCCAGGAGCTCCCGCAGGTTCGCCCGCTGCCCGCAGCGC 1210  
 Qy 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1211 TACTGGCAATGGCGCGCTTCTGAGCTGCTTGGGAACCCAGCGCGAGTGCCTCTAC 1270  
 Qy 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTGCG-----GTCAACCCAGCA 1321  
 Qy 431 Leu-----Asn 432  
 Db 1322 GCGGTGTCTGTGCGCGGAGAACCCCGAGGCTCTGTGGCGCGCGCGCGAGGAGGAC 1381  
 Qy 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1382 ACAGACCCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1441  
 Qy 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1442 GCGTGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
 Qy 473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492



1502 AACGAACGCCGCTTCTCTCAGGAACCAAGAAAGTTTCATCTCCCTCGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuLeuMetTrpTyrMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGAGGAGTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleuAla 532  
1622 AGCCAGGGGTGGCTGTGTCTCGGCGCGCAGAGCACCGTCTCGGTGAGAGATCCCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGCTGATGAGTGTAGTGTACGTCTGAGCTGCTCAGGTCTTCTTTTAT 1741  
553 IleThrCluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGAGACCAAGTTCACAAAGAACAGCGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAGAGCATGGATACACAGACACTTGAAGAGGGTGCAGCTGGGGAGGTGTGCG 1861  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGGTTCAGGACATCGGAAGAGCGAGCGCGCCCTCTGACGCTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTATATCCCAAGCTGACGGGCTGCGGGCTGCGGGCATTTGTGACATGACATGCTGTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGCGAGAGAAAGAGCGGCGAGCGCTCTCACTCGAGGGGTGAAGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGGGCGCGCGCGCTCTCTCGGGCGCTCTGTGTGGCG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACCATATCCACAGGGCTTGGCGCACCTTCTGCTGCTGGGTGGCGGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGACTTGTGTAAGTGGATGTGACGGGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCAGAAAGCGCCCATCGGCGACGTCGCGAAGCCCTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGTCTCTACCTTGACAGACCTCCAGCCGTATCGGACAGTTCGTGGCTCACCTGTGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGAGTGCCTGCTCATCGAGCAGAGCTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCCAGCAGTGGCTCTTCGACGTCTTCACGTCTTCATGGCCACACCGCCCGGTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTACGTCCAGTGCAGGGATCCGCGAGGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACCGCTGTCTACGGCGACATGAGAACAGCTGTTTTCGGGGATTCGGCGG 2632

853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTCGCTTGGTGATGATTCTTGTGGTGACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAGAACTTCTCTCAGAGCCCTGGTCGAGGTGTCTCTGAGTATGGCTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGAAGACAGTGTGTACTTCCCTGTAGAACAGAGGCCCTGGGTGGCACGGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCCAAGCGCTATTCCCTGGTGGCGCTGTCTGTCTGTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGTCTCAC 2932  
953 PheGlnSerValPheLysAlaLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGGCTTCAAGCTGGGAGGAACATGGCTGCGAAACTCTTTTGGGCTCTTCGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCGCTGTTCTGGATTGAGGTGAACAGCTCCAGCGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAATCTCTCTGTCGAGCGTACAGGTTTACGCACTGTGTGCTGCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGCGTCATCTCTGACAGC 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCTCCTCTGTCTACTCCATCTCTGAAGCCAGAACCGAGGAGTGTGCTGGGGGCCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGCGCTCTGCTCCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGTGTACCTACGTGCCACTCTCTGGGTCACTCAGGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3353 CAGACGAGCTGAGTCGGAAGCTCCCGGGAGCGAGCTGACTGCCCTGGAGGCCGAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGGCACTCCCTCAGACTTCAAGACCATCTCTGGAC 3451

RESULT 12  
AR226390  
LOCUS 4015 bp mRNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1 from patent US 644650.  
ACCESSION AR226390  
VERSION AR226390.1 GI:27264905  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclonified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Cech,T.R., Lingner,J., Nakamura,T., Chapman,K.B., Morin,G.B.,  
Harley,C.B. and Andrews W.H.  
TITLE Antisense compositions for detecting and inhibiting telomerase  
reverse transcriptase  
JOURNAL Patent: US 644650-A 1 03-SEP-2002;

```

ATUES      Location/Qualifiers
Source      1..4015
            /organism="unknown"
3E COUNT    663 a 1363 c 1275 g 714 t
IGIN
ignment Scores:
ad. No.:    4,54e-174      Length:      4015
pre:        3505.00       Matches:    719
cent Similarity: 72.94%    Conservative: 122
at Local Similarity: 62.36% Mismatches: 260
ary Match:   59.40%      Indels:     52
:            6           Gaps:      13

-09-042-460-2 (1-1122) x AR226390 (1-4015)
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115
21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGCTGCCCTGCCACAGTGTGCGCGCCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
176 CGCGGGAGCCCGCGCTTTCGCGCGCTGTGTGGCCAGTGCCTGGTGTGCTGCCCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
236 GACGACGCGCGCGCGCGCGCGCTTTCCTTCCCGCAGGTGCTCTGCCGAAAGGAGCTG 295
81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGGCCCGAGTCTGCAGAGGCTGTGCGAGCGCGCGCGAGAACGTCGTGGCTTCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGAGACGGGGCGCGCGGGGGCGCGCGCGAGGCTTTCACCCAGCAGCTGCGC 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
415 AGCTACCTGCCCAACACGGTGAACCGACGTCGTGGGGAGCGGGGCGTGGGGGCTGCTG 475
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGCGCGCGGTGGCGGACGACGTGTGCTTCACTGTGTCGTCGTCGCGCTCTTTGTG 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGGTGGCTTCCAGCTGGCGCTTACAGGTGTGCGGGCGCGCGCTGTACCACTGCGCGCT 595
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCCACTCAGCGCGCGCGCGCGCACACGCTAGT---GGACCCCGAAGCGCTCTGGA--- 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
650 -----TGCAGACGGGCTTGGACCATACGTCGAGGAGCGCGCGCTGTACCACTGCGCGCT 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
689 GTCCCTCTGGGCTTCCAGCCCGGTGCGAGGAGCGCGCGGGGCGAGTGCACGCCGAGT 748
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro--- 259
749 CTGCGCTTCCCAAGAGGCCCGCGCGCTGCGCGCTGAGCGGAGCGGAGCGCGCGT 808
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
809 GGGCAGGGGTCTGGGCGCCACCGGGGAGGCGGTGGACCGAGTGACGGTGGTTCGT 868
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292

```

1922 TTCAATCCCAAGCCCTGACGGGCTGCGGCCGATTGTGAACATGGAATACGTGCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgAlaGlyGlnAlaGlnHisPheThrGlnArgLeuLeuThrLeuPhe 652  
1982 AGACGCTCCGACAGAAAGAGGCGCCAGCGTCTCACTCGAGGGTGAAGGCACATGTTT 2041  
653 SerMetLeuAsnTyrGluArgThrIleHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCACTACGAGCGCGCGCGCCCGCCCTCTCTGGCGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCAGGCGCTGGCGACCTTCGTGCTGCTGCTGCGGGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValIysAlaPheValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGTACTTGTCAAGGTGATGTACGGCGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAlaMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGCGAGTCTATCCGACGATCATCAAA---CCCCAGACATGCTGCGTG 2278  
733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGGTATGCCGTGCTCAGAAGCGCGCCATGGGCACGTCCGCAAGGCGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGGCTCTACCTTGACAGACCTTCAGCGGTATGACGCGGTATGCGACAGTTCGTGTC 2398  
773 AspSerAspAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC---AGCCGCTGAGGATGCGGTGTCGTCGACGAGAGTCTCTCTCTGAAT 2452  
793 GluSerSerSerLeuPhePhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGGTGGCTCTTCGACGCTTCTTCACTACGCTTCATGTGCCACACGCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTCTCAGGACCTTGTGCGAGGGATCCCGCAGGCTTCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACGCTGTGCTACGCGACATGAGAAACAAGCTGTGTCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTGCGTGTGGTGGATGATTTCTTGTGGTACACCTCACCCTCACCAC 2692  
873 AlaIysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTTGTGCGAGGTGCTCCCTGAGTATGGTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGTGTAATCTTCCCTGTAGAGACGAGCGCCCTGGGTGGACGGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACGCGCTATTCTCCCTGCTGGTGGCGCTGCTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGACAGCGCACTACTCCAGCTATGCGGACCTCCATCAGAGCCAGTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGGCTTCAAGGCTGGAGGAACTGCGTGCAGAACTCTTTGGGGTCTTGGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTCAACGCTGTTTCTGGATTGCGAGGTGAACAGAGCTCCAGACGCGTGTGCACC 3052

993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAGATCTCTCTGTCGAGCGGTACAGGTTTTCAGCATGTGTGCTGAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTATCAGCAAGTTTGGAGAAGCCCAACATTTTCTCGCGCTCATCTCTGACAGC 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCATCTCTGAAAGCAAGACGAGGGATGTGCTGGGGGCGCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGCGCTCTGCTCCCTCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAAGCTGACTCGACACCGGTGCACCTAGCTGCCACTCTCTGGGGTCACTCAGGACGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3353 CAGACGCGAGCTGAGTGGAAAGCTCCCGGGGACGAGCTGACTCCCTCGAGGCGCGAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCGACTGCCCTCAGACTTCAGACCATCTCTGGAC 3451  
RESULT 13  
AR243328  
LOCUS 4015 bp mRNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1 from patent US 6475789.  
ACCESSION AR243328  
VERSION AR243328.1 GI:27290539  
KEYWORDS Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Cech,T.R., Lingner,J., Nakamura,T., Chapman,K.B., Morin,G.B.,  
Harley,C.B. and Andrews,W.H.  
TITLE Human telomerase catalytic subunit: diagnostic and therapeutic  
methods  
JOURNAL US 6475789-A 1 05-NOV-2002;  
FEATURES Location/Qualifiers  
source 1..4015  
/organism="unknown"  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,54e-174 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13  
US-09-042-460-2 (1-1122) x AR243328 (1-4015)  
993 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
993 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGCGCAGCTGCTGCGCGCGCTGCGGGGCCCGGCGCTGGTGGTGCAG 175  
993 41 ProGlyAspProLysIleTyrArgThrValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGGACCCCGCGGCTTTCGCGCGCTGGTGGCCAGTGGCTGGTGGTGGCTGG 235

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
236 GACGACGGCGCCCGCCCGCCCTCTTCCGACAGGTGCTCTCCCTGAAGAGCTG 295  
81 ValAlaValValGlnArgLeuCysGluArgGlnValValValValValPheGly 100  
296 GTGGCCCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGAAGAACGTCGTGGCTTCGGC 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerSerValArg 120  
356 TTCGGCGCTGTGAGCGGGCGCCCGCGGGCCCGCCCGAGGCGCTTCACCAACAGCGTGGC 415  
121 SerTyrLeuProAsnThrValIleGlnThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
416 AGCTACCTGCCCAACACAGGTGACCGAGCACTGCGGGGGAGCGGGGCTGGGGCTGCTG 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGGCCCGCGTGGCGAGCGAGCTGTGCTTACCTGTGTCAGCTGCGGCGCTCTTTGTG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGCTGGCTACACAGTGTGCGGGCGCGCGCTGTACACAGCTCGGGCT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCGCGCGCCCGCCACACAGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
650 -----TGCAGACCGGCGCTGGAACCCATACGTCAGGAGGCGCGG 698  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
689 GTCCCGCTGGGCTCCAGCCCGCGTCCGAGGAGCGCGGGCGAGTGCAGCGCGAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
749 CTGGCGTGTCCCAAGAGCCCGCGCTGCGCTGCGCTGAGCGGAGCGGAGCGCGGTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGCTCTGGGCGCCACCGCGGAGGAGCGGTGACCGAGTACCTGGTTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTCACTTGCAGAG-----CCCGCCGAAGAGCCACTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GGTGGCTCTCTGGCAGCGGCACTCCACCCATCCGTGGGCGCCAGCACACAGCGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCATCCGCGGCCACCACTCCCTGGGACACCGCTTGTCCCGGGGTAGCGC 1033  
331 GluThrArgHisPheLeuTyrSerArgLysPheGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACAGAGCACTTCTCTACTCTCAGCGAC---AAGGAGCAGCTGCGGCGCTCTCTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCAGGCCCGACGCTGCTGGGCTCGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCGCTGAGTCCAGGAGCTCCCGCAGGTTGCCCCCGCTGCCCGCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTGTCTGTGAGTGTCTGGAAACACAGCGGAGTGCCTTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430

1271 GGGGTGTCTCTCAAGAGCACTGCGCGCTGCGAGCTGGG-----GTCACCCAGCA 1321  
431 Leu-----Asn 432  
1322 CCCGGTGTCTGTGCCCCGGGAGAACCCAGGGCTCTGTGTGGCGCCCGCCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGTGTCTGTGTGAGCTGTCTCCGCGACACAGAGCCCTTGGCAGGTATAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGTGCGGGCTGCTGCGCGCTGTGTGCGCGGTGTGTGCGCGCTCTGGGGCTCCAGGAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492  
1502 AACGAACCGCGCTCTTCAGGAACACCAAGAGTTTATCTCTCCCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGACGAGCTGTACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCACAGGGTGTGCTGTCTGCGCGCGCAGAGCACCGTGTGCTGAGGAGATCCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTCTCTGCACTGCTGT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTCCAGGAGACCACTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTGTCAAGCATTTGAATACAGACACTTGAAGGGTGTGAGCTGCGGAGCTGTG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGTTCAGGAGCATCGGAAGACCGAGCGCGCTGTGTGACGTCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCTATCCCAAGCTCAGCGGCTGCGCGCGATTTGTGAACATGACTACGTCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGACGTTCCGACAGAAAGAGGCGCGAGCTCTCACTCGAGGGGTGAGGCACCTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACGAGCGGCGCGCGCGCTCTCTGCGCGCTCTCTGCTGTGTGCGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATTCACAGGCGCTGGCGACCTTCTGCTGCTGTGCGGCGCCAGGACCGC 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGGTGGATGTGCGGCGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTATCGCCAGCATCATCAA---CCCCAGAACACGACTGCTGCTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGGTATCGGTGTGTCAGAGCGCGCCCGCCAGCTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772

```
2339 CAGCTCTCTACCTTGACAGACCTCCAGCGGTATACGAGACGTTCTGGTCCACCTGCAG 2398
773 AspSerAspAlaSerAlaLeuArgAenSerValValIleGluInSerIleSerMetAen 792
      :::::
2399 GAGACC-----AGCCCGCTGAGGATCGCTCGTATCAGCAGAGCTCTCCCTGAAT 2452
793 GluSerSerSerLeuPhePhePheLeuHisPheLeuArgHisSerValValIys 812
      :::::
2453 GAGCCAGAGTGGCTCTTCGAGCTTCTTCAGCTTATGTGCCACAGCCGCTGGCG 2512
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
      :::::
2513 ATCAGGGCAAGTCCCTAGTCCAGTCCAGGGGATCCCGACGGTCTCCATCTCTCCACG 2572
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAenLysLeuPheAlaGluValGlnArg 852
      :::::
2573 CTGCTCTGCAGCTGTGCTACGGCGCATGGAGAACAAAGCTGTTTTCGGGGATTCGGCGG 2632
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2633 GACGGCTGCTCTCGCTTGTGGTGGATGATTTCTTGTGTGACACCTCACCTACCCAC 2692
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAen 892
2693 GCGAAAACCTTCTCAGGACCCCTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGTGAAC 2752
893 LeuGlnLysThrValValAenPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
      :::::
2753 TTGGGAGACAGTGTGAACTTCTCTGTAGAGACGAGGCCCTGGGTGGACGGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATGCGCGCCACCGCCCTATTCCTCTGTGGCGCTGCTGTGATACCCGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGGAGGTGACAGCGACTACTCCAGTATGCCGAGACCTCCATCAGACCGACTCTCACC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2933 TTCAACCGCGGCTCAAGGCTGGAGGACATGCTGCTCGCAACTCTTTGGGTCTTGGG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGCCCTTCTCGATTTTCAGGTGAACAGCTCCACAGCGGTGTGCACC 3052
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAAGATCTCTCTGCTGAGCGGTACAGGTTTCACGCATGTGTCTGCAGCTC 3112
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTTCATCAGCAAGTTTGGAGAACCCCACTATTTTCTCGCGTCTATCTCTGACACG 3172
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCCTCTCTGTACTCTCATCTCTGAAGCCAGACGAGGATGTGCTGGGGCCCAAG 3232
1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGGCCCGCGCCCTCTGCTCTCCAGCGCGTGTGAGTGTGTGCTGGGATTCCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAGCTGACTGACACCGTGTACCTAGTGCCTCCTGGGTCTACTCAGGACGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
3353 CAGACGAGCTGAGTCGAAGCTCCCGGGACGACGCTGACTGCTGAGCGCGCAGCC 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCCGCACTGCCCCCTCAGACTTCAAGACCATCTCTGGAC 3451
```

## RESULT 14

```
AR263555 LOCUS 4015 bp DNA linear PAT 29-JAN-2003
DEFINITION Sequence 3 from patent US 6331399.
ACCESSION AR263555
VERSION AR263555.1 GI:26075300
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4015)
AUTHORS Monia,B.P., Gaarde,W.A. and Wanciewicz,E.
TITLE Antisense inhibition of tert expression
JOURNAL Patent: US 6331399-A 3 18-DEC-2001;
FEATURES Location/Qualifiers
     source          1..4015
                     /organism="unknown"
```

BASE COUNT 663 a 1363 c 1275 g 714 t

## ORIGIN

## Alignment Scores:

Pred. No.: 4.54e-174 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13

US-09-042-460-2 (1-1122) x AR263555 (1-4015)

```
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGCGCGCTGCGCGCTGCTGCTGCGCGCGCTGCGCGAG 115
QY 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGAG 175
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
Db 176 CGCGGGACCGCGCGCTTCCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCTGCG 235
QY 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
Db 236 GAGCAGCGCGCGCGCGCGCGCGCTTCCCTTCCGCGCGCTGCTGCTGCTGCTGCTGCTG 295
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
Db 296 GTGGCCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCTGCG 355
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
Db 356 TTGCGGCTGTGCGAGCGCGCGCGCGCGCGCGCTTCCCTTCCGCGCGCTGCTGCTGCTG 415
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
Db 416 AGTACCTGCCCCAACACCGTGTGCGAGCGCTGCGCGCGCGCGCGCGCGCGCGCTGCTG 475
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
Db 476 GTGGCGCGGTGGCGAGCAGCTGTGCTGCTTCACTGCTGCGACCGCTGCGCGCTGCTG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
Db 536 CTGGTGGCTGCCAGCTGCGCTTACAGGTGTGCGCGCGCGCGCGCTGTACAGCTGCGCG 595
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
Db 596 GCATCTCAGCGCGCGCGCGCGCGCGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGluAlaPro 220
      :::::
      :::::
      :::::
```

```
550 -----TCCGAACGGGCTCGAACCATAGCGTCAGGAGCGCCGG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
689 GTCCCCCTGGCCCTGTCAGCCCGGGTGGAGAGCGCGGGGCGAGTCCAGCGCAAGT 748
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluPro--- 259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
749 CTGCCGTTCGCCAAGAGCCAGCGCTGGCGCTGCCCTTACGCCGAGCGAGCCCGTT 808
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
809 GGGCAGGGGCTCTGGGCCACCCCGGCGAGAGCGCGTGGACCGAGTACCGTGTTCGT 868
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
869 GTGTGTACCTCCAG-----CCGCCGAAGAAGCCACCTCTTTTGGAG 913
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
914 GTTGGCTCTCTGCAGCGGCCACTCCACCCATCGTGGCGCGCCAGCACCGCGGGC 973
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
974 CCCCCATCCATCGCGGCCACACACTCCCTCGGACACGCCCTTGTCCCGGTGTACGCC 1033
331 GlyThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1034 GAGACCAAGCACTTCTTACTCTCTCAGGGAC---AAGAGCAGTGGCGGCTCTTTC 1090
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1091 CTACTCAGCTCTCGAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGAGACCATCTTT 1150
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1151 CTGGTTCAGAGCCCTGGATGCCAGGAGATCCCCGAGGCTTCCCGCGCTGCCCGAGCGC 1210
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1211 TACTGGCAATGCGGCCCTGTCTTCGGAGCTGCTTGGGAACCAACAGCGCAGTGCCTTAC 1270
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1271 GGCGTGTCTCTAAGACGCACTGCCCGCTGCGAGCTGGG-----GTCAACCCCAACA 1321
431 Leu-----Asn 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1322 GCCGCTCTGTCCCCGGGAGAGCCCCAGGCTCTGTGGCGCCCCCGAGGAGGAGC 1381
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1382 ACAGACCCCGCTGCTGTGTGACAGCTGCTCCGCGACAGCAGCAGCCCTTGGCAGGTGAC 1441
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1442 GGCTTGTGGCGGCTGCTTGCCTGCCCGCGCTGGTGGCCCGAGGCTCTGGGGCTCCAGGAC 1501
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1502 AACGAACCGCGCTCTCTCAGGAACCAAGAGTTTCATCTCTCTCTGGGAAGCATGCCAAG 1561
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1562 CTCTCGCTCAGAGGTGACGTGGGAAGATGACGCTGGCGGACTTGGCTTGGCTGGCAGG 1621
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1622 AGCCAGGGGTTGGCTGTCTTCGGCGCGCAGAGCACCGCTCTCGTGTAGAGAGATCCTGGCC 1681
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1682 AAGTTCTCTGCACTGGTGTAGTGTACGTCTGTCGAGCTGCTCAGGTCTTCTTTTAT 1741
```

```
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1742 GTCACGGAGACACACGCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1802 AAGTTGCAAGCATTTGGATCAGACAGCACTTGAAGAGGTGACAGCTGGGGAGCTGTCC 1861
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1862 GAAGCAGAGGTTCAGCAGCATCGGAGAGCCAGGCGCCCTGTCTGAGCTCCAGACTCCGC 1921
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1922 TTCAATCCCAAGCTTGGGCTGGCGCGGATTTGAACATGGACTAGCTGCTGGAGACC 1981
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1982 AGAACGTTCCGACAGAAAAAGAGGCGCGGTCTCACCTCGAGGGTGAAGGCACTGTTC 2041
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2042 AGCTGTCTCACTACGAGCGCGCGCGCCCGGCTCTCTGGCGGCTCTGTCTGCTGGC 2101
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2102 CTGGACGATATCCACAGGCGCTGGCGCACCTTCGTCTGCTGTGGCGGCCAGGACCCG 2161
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2162 CGCCTCTAGCTTACTTTGTCAAGTGTGATGTGACGGCGGTACGACACCATCCCCAG 2221
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2222 GACAGCTCAGGAGTCTATCCGACGATCATCAA---CCCCAGAACACATCTGCTGGTG 2278
733 ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2279 CGTGGTATCGCTGTCAGAAAGCGCCCATGGCACGTCGCGAAGGCTTCAAGAGC 2338
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2339 CACGCTCTTACCTTACAGACTCCAGCGGTACATGGACAGTTCGTGGCTCAGCTCAG 2398
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2399 GAGACC-----AGCCCGCTGAGGAGTCCGCTCTCATCGAGCAGACTCCTCCTGAAT 2452
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValLys 812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2453 GAGGCGCAGTGGCTCTTCGACGCTTCTTACGCTTCATGTGCCACCCAGCGCGTCCG 2512
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2513 ATCAGGGCAAGTCTTACGTCAGTGCAGGGGATCCGCGAGGCTCCATCCTCTCCAG 2572
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2573 CTGCTCTGCGAGCTGTGTACGGGACATGAGAAACAAGCTGTTTGGGGGATTCGGCGG 2632
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2633 GACGGCTGCTCTCTCGTTTGGTGATGATTTCTTGTGTGACACCTTCACTCACCAC 2692
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2693 GCGAAAACTTCTCTCAGGACCTTGTCCGAGGTGTCTCTGAGTATGGCTGGTGGTGAAC 2752
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2753 TTGCGGAACAGACGTGGTGAACCTTCCCTGTAGAACAGAGGCCCTGGGTGGCACGGCTTT 2812
```





351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTGAGGCGCAGCTGACTGCTCGGAGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTCACAGGCTCGATGTCACAGGACTCCCGCAGGTTGCCCGCTCGCCAGGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGCAATGCGGCCCTGTTTCTGAGAGTCTGTTGGGAACACACGCGCTGCCCTTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGTGTCTCTCAAGACGACTGCCCGCTGCGAGTGGG-----GTACCCCGACA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTGCCGGGAGAACCCAGGCTCTGTGGCGGCCCGCAGGAGGAGAC 1381  
433 ThrSerProHisLeuMetAspLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTCGCTGTGGAGCTGCTCCGCGACAGCAGGCGCTCGGAGGTGTAC 1441  
453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGCTTGTGGGGCTCGCTGCGCGGCTGGTGGCCCGAGGCTCTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPheLeuLeuLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AAGCAACGCGCTCTCTCAGGACACCAAGAGTTCATCTCCCTCGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGTCAGAGGTGACGTGGAGATGAGCGTGGCGGACTGGCTTGGTCGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGGTGGCTGTGTTCCGGCGCAGAGCACCGCTCGCTGAGGAGATCCGCGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuArgSerPhePheTyr 552  
1682 AGTTCTCGACCTGGCTGATGAGTGTGACGTGCTGAGCTGCTCAGGTCTTCTTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGGAGACCAAGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGACAGTCCGGGAGCTGTCG 1861  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACAGAGTTCAGGACGATCCGGAAGCAGCGCCGCTCTGCTGACGTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTCAGCGGCTGCGCGCGATGTGAACATGAGACTACGTGTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGTTCGCGAGAAAGAGCGCGGCGCTCTCACTCCAGGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCTGGCGGCTCTGTGTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACATATCCACAGGCGCTGGCGACCTTCTGCTGCTGCTGCTGCGGCGCCAGACCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712

2162 CGCCTCAGCTGTACTTTGTCAAGGTGATGTGAGGCGGTACGACACATCCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGTTCATCCAGCATCATCAAA--CCCCAGAAACAGTACTGGGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTCCGTATGCGGTGTCAGAGCGCCCATGGCAGCTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGCTCTACTCTGACAGACTCCAGCGGTACATGCGACAGTTCGTGGGTCACTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCCGTCTCATCGACAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGACAGTGGCTCTTCGAGCTTCTCTAGCTTCTATGTGCCACACGCGGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCGCAGTCTAGCTCAGTGCAGGCGATCCGCGAGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluLeuLysLeuPheAlaGluValArg 852  
2573 CTGCTCTGAGCTGTGTACGGGACATGAGAGAACAGCTGTTCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCGTTGGTGGATGATTTCTTGTGTGACACCTCACTCACTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GGGAAACCTTCTCAGACCTCGTCCGAGGTGCTCCCTGAGTATGGCTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAGACAGTGGTGGACTTCCCTGTAGAACGAGGCCCTGGGTGGCAGCGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACGGCTATTCCTCTGGTGGCGCTGCTGCTGGATACCCGAGC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGACAGGCGACTTCTCCAGCTATGCCCGGACCTCCATCAGAGCGAGTCTCAC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGGAGGAACATGCTGCGAACTCTTTGGGTCTTCGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGGCTCTTCTGAGATTGAGGTGAACAGCTCCAGAGCGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGTCAGGCGGTACAGGTTTACGCAATGTGTGTGCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCTATCAGCAAGTTTGAAGAACCCCATTTTCTCTCGCGCTCATCTCTGACACG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCTTCCTCTGCTACTCATCTGAAAGCCAGAACCGAGGATGTGCTGGGGCCAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

n on: February 1, 2004, 15:05:58 ; Search time 663 Seconds  
(without alignments)  
4568.280 Million cell updates/sec

file: US-09-042-460-2  
irect score: 5901  
quence: 1 MTRAPCPAVRSILRSRYRE.....TILKAADPALSTDFQTILD 1122

oring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

arched: 2552756 seqs, 1349719017 residues

tal number of hits satisfying chosen parameters: 5105512

nium DB seq length: 0  
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

mmand line parameters:

ODEL=frame+ p2n.model -DEV=xlh  
=/cgn2 1/USPTO.spool/US09042460/runat 01022004.150550 17004/app query.fasta\_1.1287  
B-N Geneseq 19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCI=0  
COPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
IST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
QDE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
SER=US09042460 -CGEN 1.1.0 @runat 01022004.150550 17004 -NCPU=6 -ICPU=3  
O MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
EV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

tabase : N Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3513	59.5	4015	20	AAZ08150 Human telomerase r
2	3505	59.4	3396	22	AAH44366 Human telomerase n
3	3505	59.4	3396	22	AAH48235 Heart muscle cell
4	3505	59.4	3396	22	AAH49601 Human coding seque
5	3505	59.4	3399	24	ABZ35720 Human telomerase r
6	3505	59.4	3399	24	ABX09963 Human telomerase r
7	3505	59.4	3399	24	ABV78144 Human telomerase r
8	3505	59.4	3399	24	ABL91685 Human polynucleoti
9	3505	59.4	3798	19	AAV27876 Human telomerase p
10	3505	59.4	3955	19	AAV22379 Human telomerase r
11	3505	59.4	4015	20	AAZ30154 cDNA encoding a hu
12	3505	59.4	4015	20	AAZ20279 Human telomerase r
13	3505	59.4	4015	20	AAZ00724 Human telomerase c
14	3505	59.4	4015	22	AAH45901 Human hTERT gene.
15	3505	59.4	4015	24	AAH46821 Human telomerase r
16	3505	59.4	4015	24	ABA97534 Cancer cell discr
17	3505	59.4	4015	25	ABZ22474 Human telomerase r
18	3505	59.4	4015	25	ABZ18391 Group iii cDNA can
19	3505	59.4	4042	20	AAV72117 Human catalytic te
20	3505	59.4	4070	24	ABL53711 Human telomerase c
21	3498	59.3	3396	20	AAH18266 Telomerase coding
22	3498	59.3	3964	20	AAH18254 Human telomerase c
23	3498	59.3	4023	19	AAV60320 Human telomerase g
24	3498	59.3	4027	20	AAH99424 Human EST2 coding
25	3498	59.3	4027	21	AAA29388 hEST2, a human tel
26	3496	59.2	3453	25	ABZ76217 Human TERT coding
27	3496	59.2	4037	19	AAV22428 Human telomerase r
28	3496	59.2	13766	24	AAH46790 pGRN145 plasmid DN
29	3473	58.9	8742	24	AAH46793 pWG5a plasmid DNA
30	3449	58.4	3500	20	AAH18275 Telomerase coding
31	3435	58.2	3918	20	AAH18269 Telomerase coding
32	3435	58.2	3918	20	AAH18278 Altered C-terminus
33	3304.5	56.0	3203	19	AAH18268 Human telomerase r
34	3276.5	55.5	3855	19	AAV22382 Altered C-terminus
35	3275.5	55.5	3323	20	AAH18277 Altered C-terminus
36	3234.5	54.8	3167	20	AAH18271 Altered C-terminus
37	3234.5	54.8	3167	20	AAH18280 Altered C-terminus
38	3141	53.2	7688	20	AAH18351 Telomerase coding
39	3038	51.5	7797	20	AAH18350 Telomerase coding
40	2881.5	48.8	3069	20	AAH18267 Truncated telomera
41	2877.5	48.8	3069	20	AAH18276 Truncated telomera
42	2814.5	47.7	3033	20	AAH18270 Truncated telomera
43	2814.5	47.7	3033	20	AAH18279 Truncated telomera
44	2808.5	47.6	7615	20	AAH18349 Telomerase coding
45	2790.5	47.3	2848	19	AAV27872 Human telomerase p

ALIGNMENTS

RESULT 1

AAZ08150  
ID AAZ08150 standard; cDNA; 4015 BP.  
XX  
XX AC  
XX AAZ08150;  
XX  
XX 17-JAN-2000 (first entry)  
DT  
DE Human telomerase reverse transcriptase cDNA.  
XX  
XX Human telomerase reverse transcriptase; hTERT; telomerase; hEST2;  
KW Catalytic protein component; cell proliferative capacity;  
KW cell immortality; neoplastic phenotype; diagnostic application;  
KW prognostic application; telomerase related condition; cancer;  
KW Therapeutic agent; telomerase expression; telomerase activity; ds.  
XX  
XX Homo sapiens.  
OS

Key:	Location/Qualifiers
CDS	56..3454
	/*tag= a
	/product= "Human telomerase reverse transcriptase"
	/transl_except= (pos:1877..1879, aa:Gln)

WO9950279-A1.

07-OCT-1999.

31-MAR-1999; 99WO-US07160.

31-MAR-1998: 98US-0052919.

(GERO-) GERON CORP.  
(UYTE-) UNIV TECHNOLOGY CORP.

Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB, Andrews WH;

WPI: 1999-610834/52.

P-PSDB; AAY28881.

Antisense polynucleotides for human telomerase reverse transcriptase used for diagnosing or treating cancer -

Claim 1; Fig 1; 31pp; English.

The present sequence encodes for human telomerase reverse transcriptase (hTERT). This is the catalytic protein component of telomerase and is also referred to as hE23. This correlates with cell proliferative capacity, cell immortality, and the development of a neoplastic phenotype. Human TERT antisense oligonucleotides are useful for diagnostic or prognostic applications to telomerase related conditions, including cancer. They are also useful as therapeutic agents, for inhibition of telomerase expression and activity.

Sequence 4015 BP; 663 A; 1364 C; 1274 G; 714 T; 0 other;

ignment Scores:	
ed. No.:	3,32e-234
ore:	3513.00
cent Similarity:	73.03%
st Local Similarity:	62.45%
ery Match:	59.53%
:	20
	29
	13
	4015
Length:	4015
Matches:	720
Conservative:	122
Mismatches:	259
Indels:	52
Gaps:	13

-09-042-460-2 (1-1122) x AAZ08150 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20  
56 ATCCGGCGCGCTCCCGCTGCCAGCGCGCGCTCCCTGCTGCGACCACTACCCGAG 115  
21 ValTrpLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
116 GTGCTGCGCGCTGCCACGTTGCTGCGCGCTGCGGCCCGCCAGCGCTGCGGTGCGAG 175  
41 ProGlyAspProLysrIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
176 CGCGGGACCCCGCGGCTTTCGCGCGCTGCGGCCAGTGCCTGGTGCCTGCGCTGG 235  
61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
236 GACGACACGGCGCGCCCGCGCGCGCCCTCTTCGCCAGGTGCTCTGCCTGAAGAGAGCTG 295  
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
296 GTGACCCGAGTGTGCAAGAGCTGTGCGAGCGCGCGCGGCGAAGACGTGCTGGCTTCGCGC 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
356 TTCGCGTGTGTGACGGGSCCGCGGGGSCCCCGGAGGCTTTCACACACGCGTGC 415





4 Differentiation; heart muscle cell; cytokine; transcription factor;  
 4 proliferation; surface antigen; heart disease; cardiomyocyte;  
 4 bone marrow; umbilical blood cell; heart muscle degeneration;  
 4 myocardial infarction; ss.

5 Homo sapiens.

6 WO200148150-A1.

7 05-JUL-2001.

8 02-NOV-2000; 2000WO-JP07741.

9 28-DEC-1999; 93JP-0372826.

10 28-FEB-2000; 2000WO-JP01148.

11 (KYOW) KYOWA HAKKO KOGYO KK.

12 Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

13 Yamada Y;

14 WPI; 2001-425655/45.

15 P-PSDB; AAB99930.

16 Cells capable of differentiating into cardiomyocytes and originating in  
 17 bone marrow or umbilical blood cells for study of cardiomyocyte  
 18 differentiation and treatment of heart disease -

19 Disclosure; Page 141-147; 187pp; Japanese.

20 The present invention describes cells originating in bone marrow or  
 21 umbilical blood cells which are capable of differentiating into  
 22 cardiomyocytes. Also described are: (1) cardiomyocytes produced by the  
 23 differentiation of the cells; (2) a method for carrying out the  
 24 differentiation into cardiomyocytes, regulated by a promotional and/or  
 25 inhibitory factor; (3) a method for the differentiation of the cells  
 26 into cell types other than cardiomyocytes; (4) drug compositions  
 27 promoting the formation of heart muscle and regeneration of heart tissue  
 28 which contain the cells; (5) a method for the production of antibodies  
 29 which recognise the cells, especially antibodies which recognise a  
 30 surface antigen on the cells; (6) a method for screening factors which  
 31 promote the proliferation of the cells; (7) a method for immortalising  
 32 the cells by expressing telomerase in them; (8) drug compositions for  
 33 the treatment of heart disease which contain the immortalised cells; and  
 34 (9) cell-free supernatant from the culture of the cells and its use in  
 35 promoting their differentiation into cardiomyocytes. The cells are used  
 36 in the treatment of diseases involving heart muscle degeneration, such  
 37 as myocardial infarction and in the study of cardiomyocyte  
 38 differentiation. AAH4351 to AAH4409 and AAB99915 to AAB99935 represent  
 39 sequences used in the exemplification of the present invention.

40 Q Sequence 3396 BP; 549 A; 1157 C; 1088 G; 602 T; 0 other;

41 Alignment Scores:

red. No.:	9,49e-234	Length:	3396
core:	3505.00	Matches:	719
Percent Similarity:	72.94%	Conservative:	122
Best Local Similarity:	62.36%	Mismatches:	260
Query Match:	59.40%	Indels:	52
B:	22	Gaps:	13

42 S-09-042-460-2 (1-1122) x AAH44366 (1-3396)

43 Y	1	MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu	20
44 b	1	ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGAGCCACTACCGGAG	60
45 Y	21	ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln	40
46 b	61	GTGTCCTCCCTGCGACGTTCTGTCGGCGCTTGGGCGCCCGAGGCTGGCGCTGTGTG	120
47 Y	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr	60

121	CGCGGGGACCGCGCGCTTCCGCGGGCTGTGTGGCCAGTGCCTGTGTGGTGCCTGG	180
61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
181	GAGCAGCGCGCGCGCGCTTCCGCGAGGTGTCTGCTGCTGAAGGAGCTG	240
81	ValAlaArgValValGlnArgCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
241	GTGCGCGAGTGTGCAGAGGCTGTGCGAGCGCGCGGAGAGACGTGTGCTGCTTGGC	300
101	PheGluLeuLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerVal	120
301	TTGCGCTGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	360
121	SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaTyrMet	140
361	AGTACTGTGCTCCCAACACGCTGACCGGACGACCTGCGGGGAGCGGGGCTGTG	420
141	LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu	160
421	CTGCGCGCTGTGCGCGACGACGCTGTGTTCACCTGCTGGCAGCGCTCTTTGTG	480
161	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCys	180
481	CTGCTGCTCCAGCTGCGCTACCAAGTGTGCGCGCGCGCGCTGTACAGCTCGCG	540
181	ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProVal	200
541	GCACCTCAGCG	594
201	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAla	220
595	-----TGCAGAACCGCGCTGGACCATAGCGTCAGGAGCGCGCG	633
221	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer	240
634	GTCCCCCTGGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	693
241	ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGly	259
694	CTGCGCTGTGCG	753
260	-----HisArgGlnValLeuProThrProSerGlyLysSerTrp	272
754	GGCGAGCGCTGCG	813
273	ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSer	292
814	GTGCTGTCACTTCCGCGA-----CCGCGCGAGAGCGCGCGCGCGCGCG	858
293	GlyLysValSerAspLeuSerLeuSer-----GlySerValCysLysHisLys	311
859	GGTGGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	918
312	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro	330
919	CCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	978
331	GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnPro	350
979	GAGACCAAGCACCTTCTCTACTCTCAGCGCGA-----AAGGAGCAGCTGCG	1035
351	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeu	370
1036	CTACTAGCTCTCTGAGGCCACGCTGCTGCGCGCTGCGCGCTGCGAGACCAT	1095
371	LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSer	390
1096	CTGGTTCCAGCGCTGAGTCCAGGCGACTCCCGCGAGGTTGCGCGCGCTGCG	1155
391	TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCys	410
1156	TACTGGCAAAATGGCGCGCTGTTTCTGGAGTCTGTGGGAGACCGCGCGCTG	1215



3358 AACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGAC 3396

SIII.T 3

H48235

AAH48235 standard; DNA: 3396 BP.

AAH48235:

21-SEP-2001 (first entry)

Heart muscle cell differentiation related DNA SEQ ID NO: 32.

Heart muscle cell; human; cell differentiation; heart disease; ds.

**Homo sapiens.**

WO200148151-A1.

05-JUL-2001.

27-DEC-2000: 2000WO-JP09323.

28-DEC-1999: 99JP-0372826.

28-FEB-2000; 2000WO-JP01148.

02-NOV-2000; 2000WO-JP07741.

(KYOW ) KYOWA HAKKO KOGYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

Yamada Y;

WPI; 2001-425656/45.

P-PSDB; AAG64859.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease -

Disclosure; Page 147-153; 183pp; Japanese.

The present invention provides cells originating in the human bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. These cells are useful in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction, and the study of cardiomyocyte differentiation. The present sequence is an oligonucleotide described in the exemplification of the invention.

Sequence 3396 BP; 549 A; 1157 C; 1088 G; 602 T; 0 other;

ignment Scores:

ed. No.:

Score: 3505.00

Percent Similarity: 72.94%

Best Local Similarity: 62.36%

every match: 59.40%

77

1-09-042-460-2 (1-1122) x AAH482

1000

3  
3  
4  
5  
4  
4  
2  
3  
7  
4  
5  
4  
4  
3  
4  
7  
4  
4  
4  
3  
3  
4

1 ATGCCGGCGCTCCCGCTGCCGAGCCGCTCCCTGCTGGCAGCCACTACCGGAG 60

21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40

130

0 1 2 3 4 5 6 7 8 9 A B C D E F

41 ProGluAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60

[illegible]

1216 GGGGTGCTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCG-----GTCAACCCAGCA 1266  
431 Leu-----Asn 432  
1267 GCGGTGCTGTGTCGCCGGAGAGCCCAAGGCTCTGTGGCGGCCCGAGAGGAGGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrTrpGlnValTyr 452  
1327 ACAGACCCCGCTGCTGTGTCAGCTGCTCCCGCAGCAGCAGCGCCCTGCGAGGTGTAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472  
1387 GGGTTCGTGGCGGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCGCGC 1446  
473 AspGluArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492  
1447 AACAAAGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAACATGCCAAG 1506  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCTCGCTGCGAGGCTGACGTGGAAGATGACGCTGCGGACTGCGCTTGGCTGGCGAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCCAAGGGTGGCTGTGTTCCGCGCGCAGAGCACCGCTCTGCGTGAGGAGATCTCGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1627 AAGTTCCTGCACTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTCAACGAGACCAAGTTCCTCAAGAACCAAGCTCTTTTCTACCGGAAGAGTGTGGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTCGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1806  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAACGAGAGGTGAGGAGCATCGGGAAGCAGGCGCGCGCTGCTGACGTCCAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1867 TTATCCCCAAGCTGACGGCTGCGCGCTGCGATGTGAACATGACATGCTGCTGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAAGCTTCGCGAGAGAAAGAGGCGCGAGCTCTCACTCGAGGGGTGAAGCACTGTTTC 1986  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 AGCGTGTCAACTACAGCGCGCGCGCGCGCGCGCGCGCGCGCTCTGCTGGCGCTGCTGGC 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCAAGAGGCTTGGCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CGCGCTGAGCTGACTTCTCAAGGTGGATGTGAGCGGCGCGCGCGCGCGCGCTCTGCTGGC 2166  
713 GlyLysLeuValGlnValAlaAlaMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACAGGCTACGGAGGTCTACCGCATCATCAAA---CCCCAGAACACATCTGCGGTG 2223  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 CGTCGCTATGCTGCTGCTCAGAAGCGCGCCATGGCAGCTTCGCGCAAGGCTTCAAGAGC 2283  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CAGCTCTCTACCTTGACAGACCTCCAGCGCGTACATGCGACAGATTCCTGGCTCACCTCGAG

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2344 GAGACC-----AGCCCGCTGAGGATGCGGTCTCATCAGCAGAGCTCTCCCTGAAT 2397  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAGCCAGCAGTGGCTCTTCAGCGTCTTCCTACGCTTCATGTGCCACACCGCGTGGCG 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCAGGGCAAGTCTCTAGCTGCTGCGGAGATCCCGCAGGCTCCATCTCTCCACG 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 CTGCTCTGAGCGCTGTGTACGGGACATGGAGAACAGCTGTGTGGGGGATTCGGCGG 2577  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2578 GACGGGCTGCTCTGCGTGTGGTGTGATGATTTCTGTGTGTGACACCTCACCTCACCCAC 2637  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 GCGAAACCTTCTCTCAGACCTGCTCGAGGTCTCCTGAGTATGCTCGGTGGTGAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGGAAGACAGTGGTGAATCTCCCTGTAGAAGACGAGCGCTGGGTGGACGGCTTTT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATGCGCGGCCACGGCTATTCCCTGCTGGTGGCGCTGTGTGGATACCCGAGC 2817  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGGAGGTGCGAGGCACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCACTCTCACC 2877  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGCGCTCAAGGCTGGGAGAACATGCTGCGCAACTCTTTTGGGGTCTTGGCG 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 CTGAAGTGTACAGCTGTTTCTGGATTTGAGTGAACAGCTCCAGAGCGGTGTGACC 2997  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATCTACAAGATCTCTCTGCTGCGGCTGACAGGTTTTCAGCATGTGTGTGCTGAGCTC 3057  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTTCTATCAGCAAGTTTGGAGAACCACCACTTTTCTGCGCGCTCATCTCTGACAGC 3117  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCTCTCTGCTACTCTCATCTCGAAGCCACAGACCGAGGATGTGCTGGGGGCCAAG 3177  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3178 GCGCGCGCGCGCTCTGCTGCGGCGCGCTGCGATGGCTGTGCCACCAAGCATTCCTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 CTCAGCTGACTCGACCCGCTGCTACCTGCTGCGCACTCTGGGCTCCTCAGACAGGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3298 CAGACGAGCTGAGTCCGGAAGCTCCCGGGAGCAGCTGACTGCTGCTGGAGCGCGCAGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCGGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3396

Qy	101	PheGluLeuLeuAsnGlnAla	ArgGlyClyProProMetAlaPheThrSerVal	Arg 120
Db	301	TTGGCGTGTGTGACGGGGCCCGGGGGCCCCCGAGGCCTTACACCACGAGTGCGC		360
Qy	121	SerTyrLeuProAsnThrVal	IleThrLeuArgValSerGlyAlaTrpMetLeuLeu	140
Db	361	AGCTACTTCCCAACACAGGTGACACGACCACTGCGGGGAGCGGGCGTGGGGCGTGC		420
Qy	141	LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu		160
Db	421	CTGGCGCGGTGGGGACAGCATGTCTGTTCACCTGTGTGGACGCTGCGCGCTCTTTGG		480
Qy	161	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla		180
Db	481	CTGGTGGCTCCCAAGCTGGCCCTTACCAGGTGTGCGGGCGCGCGCTGTACCAAGCTCGCGGT		540
Qy	181	ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg		200
Db	541	GCCACTCAGGCCCGGGCCCCCGGCCACACCTAGT---GGACCCCGAAGCGCTCTGGGA---		594
Qy	201	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro		220
Db	595	-----TGCGAACGGCGCTTGGAAACCATAGCGTCAGGAGGAGCGCG 633		
Qy	221	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer		240
Db	634	GTCCCCGTGGCGCTGCACGCCCGGTGCGAGAGCGCGGGGCGAGTGCACGCCGAAGT		693
Qy	241	ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro---		259
Db	694	CTGCGTGTGCCAAGAGGCCACAGCGGTGCGCTGCCCTGAGCGGAGCGACGCCGTT		753
Qy	260	-----HisArgGlnValLeuProThrProSerGlyLysSerTrp---		272
Db	754	GGGACGGGGTCTTGGGCCACACCGGGCGAGACGGGTGACCGAGTACCGTGTCTGTGT		813
Qy	273	ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys		292
Db	814	GTGTGTCACTCCACAGA-----CCGCGCGAGAGCACCTCTTTGGAG 858		
Qy	293	GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer		311
Db	859	GGTGCGTCTCTGCACGGCGGCACATCCACCATCCGTGGCGCCGCCACACACGCGGGC		918
Qy	312	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---		330
Db	919	CCCCATCCACATCGGGGCCACACAGTCCCTGGGACGCCCTTGTCCCGGTGTACGCC		978
Qy	331	GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe		350
Db	979	GAGACCAAGCACTTCTCTACTCTCCACGGCAC---AAGGAGCAGCTGCGCGCTCTCTC		1035
Qy	351	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe		370
Db	1036	CTACTCAGCTCTCTGAGGGCCAGCTGACTGCGGTGCGAGGCTGTGGAGACCATCTTT		1095
Qy	371	LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg		390
Db	1096	CTGGGTTCCAGGCCCTTGGATGCCAGGACTCCCGCAGGTGTGCCCGCGCTGCCCGCAGC		1155
Qy	391	TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr		410
Db	1156	TACTTGGCAATGGCGGCCCTGTCTTCGTGAGCTGCTTGGGAACCAACGCGAGTGCCTAC		1215
Qy	411	ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla		430
Db	1216	GGGTGTCTCTCAGACGCACTGCCGCTGCGAGTGGC-----GTCAACCCACGA 1566		
Qy	431	Leu-----		432
Db	1267	GCGCGTGTCTGTGCCCGGAGAGCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAC		1326

```

433 ThrSerProHisLeuMetAspLeuArgLeuHisSerSerProTrrpGlnValTyr 452
1327 ACAGACCCCGTCCGCTGGTGGAGCTCCGCCAGACAGACAGCCCTCGGAGGTATC 1386
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrrpGlnHis 472
1387 GGCTTCGTGGGGCTCGCTCGCGCTGGTGGCCCGCCAGGCTCTGGGGCTCCAGGCAC 1446
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
1447 AACGAACCGCGCTTCCTCAGGAACACCAAGAGTTTCATCTCCCTGGGAAGCATCCCAAG 1506
493 LeuSerLeuGlnGluLeuMetTrrpLysMetLysValGluAspCysHisTrrpLeuArgSer 512
1507 CTCTCGCTGAGAGCTGAGCTGGAAGTGGCTGGGAGCTGGGCTGGCTGGCGAGG 1566
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgGluLeuAla 532
1567 AGCCACAGGGTGGCTGTTCGGCGCCGAGACACCGTCTCGCTGAGGAGATCTCGGCC 1626
533 ThrPheLeuPheTrrpLeuMetAspTrrpValValGluLeuLeuArgSerPhePheTyr 552
1627 AAGTTCCTGACCTGGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrrpSer 572
1687 GTACAGGAGACACAGTTTCAAGAACAGAGCTCTTTTCTACCGGAAGATGCTCGGAGC 1746
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1747 AAGTTCGAAGCATTTGAATCAGACAGCATTTGAAGAGGGTGCAGCTCGGGAGTGTGCG 1806
593 GlnGluGluValArgHisGlnAspTrrpLeuAlaMetProIleCysArgLeuArg 612
1807 GAACAGAGTTCAGGAGCATTCGGAGCCAGCCCGCTGCTGAGCTCCAGATCTCCGC 1866
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1867 TTCATCCCAAGCTTGAGGGCTCGCGCCGATTTGTAACATGAGCTACGCTCGGGAGCC 1926
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1927 AGAAGTTCGACAGAAAGAGGGCCGAGCTCTCACCTCGAGGGTGAAGCATCTGTTTC 1986
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
1987 AGCTGTCAACTACGAGCGCGCGCGCCGCTCTCTGGCGCCCTCTGTGTGGGC 2046
673 MetAsnAspIleTyrArgThrTrrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2047 CTGACCATATCCACAGGGCTGGCGCACCTTCTGCTGCTGCTGGGGCCCGAGACCGG 2106
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712
2107 CCGCTGAGCTGACTTTGTCAGGTGTGTGACGGGGCGGTACGACACCATCCCCAG 2166
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2167 CACAGGCTCAGGAGGTCTATCGCCAGCATCATCAA---CCCCAGAACACGTACTCGCTG 2223
733 ArgGlnTrrpAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2224 CGTCGGTATCCGTCGTCGAAGGGCCGATGCGACGCTCGCTGCTGGGGCCCGAGACCG 2283
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2284 CAGGTCTCTACCTTGACAGACCTCCAGCCGCTACATGCGACAGTCTGCTGCTCACCTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2344 GAGACC-----AGCCCGCTGAGGAGTGCCTGCTCATCGACAGAGCTCCCTCCCTGAAT 2397
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812

```

```

2398 GAGGCGAGAGTGGGCTCTTTCAGAGCTCTTCTAGCTTCATGTCGCCACCGCGGTGGC 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2458 ATCAGGGGCAAGTCTAGCTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2518 CTGCTCTGACGCTGTGCTACGCGCATGAGAACAGCTGTTTGGGGATTCGGCGG 2577
853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2578 GACGGGCTCTCTCTCGTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2638 GCGAAACCTTCTCAGGACCTGTGTCGAGGTGCTCTGAGTATGCTGCTGCTGGTGAAC 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyValAlaAlaPro 912
2698 TTGCGGAACAGATGGTGAACCTTCTCTGTAGAACGAGGCGCTGGGTGGCAGCGCTTT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTrrpCysGlyLeuLeuLeuAspThrGlnThr 932
2758 GTTCAGATCCGCGCCACAGGCTTATCCCTGCTGGTGGCGCTGCTGCTGATACCGGACC 2817
933 LeuGluValPheCysAspTyrSerGlyTrrpAlaGlnThrSerIleLysThrSerLeuThr 952
2818 CTGAGGTGCAGAGGACTTCTCCAGCTATGCCGAGACTTCATCAGAGCCAGTCTCACC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2878 TTCAACCGCGGCTTCAAGGCTGGAGGAACATGCTCGCAAACTCTTTGGGGTCTTTCGG 2937
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2938 CTGAAGTGTACAGAGCTGTCTGTGATTTGAGGTGAACAGCTCCAGACGGTGTGCACC 2997
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
2998 AACATCTACAAGATCTCTCTGCTGAGGCGTACAGTTTCAGCATGTGTGCTGCAGCTC 3057
1013 ProPheAspGluArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3058 CCATTTTCATCAGCAAGTTTGAAGAACCACCATTTTCTCGCGCTCATCTCTGCACAG 3117
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3118 GCCTCCTCTGTGTTACTTCATCTGAAAGCCAGAACGAGGATGTCGCTGGGGGCCAAG 3177
1053 GlySer-----PheProProGluAlaAlaHisTrrpLeuCysTyrGlnAlaPheLeu 1069
3178 GCGCGCGCGGCTCTGCTGCTCCGAGGCGCTGAGTGGCTGTGCTGCTGCTGCTGCTGCTG 3237
1070 LeuLysLeuAlaAlaHisSerValIleTrrpLysCysLeuLeuGlyProLeuArgThrAla 1089
3238 CTCAAGCTGACTGACAGCGTGTACCTGAGTGGCCACTCTCTGGGGTCACTCAGAGACCC 3297
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
3298 CAGACCGAGCTGAGTCGAAGCTCCCGGGAGCAGCGTGAAGTCTGCTGCTGGAGCGCGC 3357
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3358 AACCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3396

```

RESULT 5

ABZ35720

ID ABZ35720 standard; DNA; 3399 BP.

XX ABZ35720;

AC ABZ35720;

XX



07-FEB-2003 (first entry)

Human polymerase reverse transcriptase polynucleotide SEQ ID NO 28.

Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide; protozoicide; gene expression; antisense; tumour; infection; plasmodium; virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus; Hepatitis C virus; human papilloma virus; gene; ds.

Homo sapiens.

DE10100588-A1.

18-JUL-2002

09-TAN-2001: 2001DE-1000588

09--JAN-2001: 2001DE-1000588

(RIBO-) RIBOPHARMA AG.

Kreutzer R. Limmer S. Rost S. Hadwiger P.

WPT: 2002-683450/74.

Inhibiting expression of target genes, useful e.g. for treating tumors, by introducing into cells two double-stranded RNAs that are complementary to the target -

Claim 13; Page 25-26; 100pp; German.

The invention relates to inhibiting expression of a target gene in a cell by introducing at least two oligonucleotides (dsRNAI and II), both with a double-stranded (ds) structure of at most 49 sequential nucleotide pairs. At least part of one strand (S1, S2) of the ds structures in each of dsRNAI and II are complementary to regions in the target gene. The method uses antisense inhibition of gene expression using double stranded RNA inhibition (RNAi). The method is particularly used to treat tumours or infections, especially by plasmodium or viruses/viroids (pathogenic on humans, animals or plants). The method provides more effective inhibition of expression than known methods using a single dsRNA, even at very low concentrations. When dsRNA has at least one unpaired nucleotide at the end, stability (and thus effective concentration in the cell) is improved and efficiency can be increased further by pretreating the cells with interferon. The present sequence is that of a target DNA of the invention.

Sequence 3399 BP: 550 A: 1157 C: 1089 G: 603 T: 0 other:

ianment Scores:

Alignment Scores:					
ed. No.:	9.5e-234	Length:	3399		
ore:	3505.00	Matches:	719		
cent Similarity:	72.94%	Conservative:	122		
st Local Similarity:	62.36%	Mismatches:	260		
ery Match:	59.40%	Indels:	52		
:	24	Gaps:	13		

1-09-042-460-2 (1-1122) X ABZ35720 (1-3399)

1 Met Thr Arg Ala Pro Arg Cys Pro Ala Val Arg Ser Leu Leu Arg Ser Arg Tyr Arg Glu 20  
1 ATGCCGGCGCTCCCGGTCCGAGCCGTGGCTCCCTGCTGGCAGCCACTACCGCGAG 60

21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
61 GTGCTGCGCTGCCACGTTCTGTCTCGGCGCCCTGGGGCCCCAGGGGTGGCGGCTGTCAG 12

41 ProGlyAspProLysIleTyrArgThrLeuValalaGlnCysLeuValCysMetHisTrp  
121 CGCGGGACCGCGGGCTTCGGGGCGTGGTGGCCAGGCCCTGGTGGTGGTGGCTGG 18

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

.....

[illegible]

431 Leu-----Asn 432  
1267 GCCGGTGTGTGCCCCGGGAGAACCCAGGGGTGTGTGGCGGCCCGCCAGAGAGAGGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1327 ACAGACCCCGTGGCTGGTGCAGCTGCTCCGCCAGCACAGACGCCCTGGCAGGTGTAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLeysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGCTTCGTGGCGGCTGCTCCGCCGCGCTGGTGGCCCCAGGCCCTCTGGGGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheLeuLeuLeuLeuLeuPheLeuLeuLeuLeuLeuLeuLeu 492  
1447 AACGACCGCGCTTCCTCAGGACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAG 1506  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGCTTGGCTGGCGAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLeuAla 532  
1567 AGCCAGGGGTGTGCTGTTCGGCGCGAGACACCGTCTGGTGGAGAGATCTCGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValGlnLeuLeuLeuArgSerPhePheTyr 552  
1627 AAGTTCCTGCACCTGGCTGATGAGTGTGTACGTGTGTGCTGAGTGTCTCAGTCTTTCTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTCCGAGGACACGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGATGTCGGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTGCAAGCAATGGATCAGACAGCACTTGAAGAGGTGCGAGCTGGCGGAGCTGTG 1806  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
1807 GAAGCAGAGGTGAGCAGCATCGGGAAGCCAGCGCCCGCTGTGAGCTGCAGCTCCAGCTCC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1867 TTCATCCCAAGCTGACGGCTCGCGCCGATGTGAACATGACATGACGTGCGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAACGTTCCGCAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTTC 1986  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 ACGGTGCTCACTACAGCGGGCGGGCGGCCCGCTCTGGGGCGCTCTGTGTGGGGC 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCCAGGGCTGCGGCACCTTCGTGTGTGTGCGGGCCAGGACCCG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CCGCTGTAGCTGATCTTTGTCAAGTGTGATGACGGCGCGGTACGACACCATCCCCCAG 2166  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACGGCTCAGGAGGTTCATCGCCAGCATCATCAAA---CCCCAGAACAGTACTGCGTG 2223  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2224 CGTCGTATGCGCTGCTCCAGGAAGGCCCGCCCATGGGCAAGCTTCGCAAGGCTTCAAGAGC 2283  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CAGCTCTCTACCTTACAGACCTCCAGCGCTACATGCGACAGTTCGTGGCTCACCTGCGAG 2343

RESULT 6

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2344 GAGACC-----AGCCCGCTGAGGATGCGCTGCTCATCGAGCAGAGCTCTCTCTGAT 2397  
793 GluSerSerSerSerLeuPhePhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAGCCAGCAGTGGCTCTTCGACGCTTCTCTACGCTTCATGTGCCACACGCCCTGGCGC 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCAGGGGCAAGTCTTCATCGTCAGTCCAGGGGATCCCGCAGGGGTCTCATCTCTCCAG 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 CTGCTCTGCAGCCTGTGTACGCGCACATGAGAGAAACAGCTGTTTGGCGGATTCGGCGG 2577  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2578 GACGGGTGCTCTCTGCTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2637  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 GCGAAACCTTCTCAGGACCCCTGTGTCGAGGTGCTCCCTGAGTATGGCTGCGTGGTGAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGGAAGACAGTGTGTAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATGGCGGCCACCGGCTATTCCTCTGGTGGCGCTGCTGCTGATACCCGGACC 2817  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGAGGTGACAGCAGCACTACTCCAGCTATGCCGAGACCTCCATCAGAGCCAGCTCTACC 2877  
953 PheGlnSerValPheValAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGCGCGCTTCAAGGCTGGAGGAACATGCTGCGCAACTCTTTGGGGTCTTGGCG 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 CTGAAGTGTACAGCCTGTTCTTCGATTTGCAAGTGAACAGAGCTCCAGACGGTGTGCAC 2997  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATCTACAGATCTCTCTGCTGAGGGGTACAGGTTTTCACGATGTGTGTGCTGAGCTC 3057  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTTTCATCAGCAAGTTTGAAGAAGACCCACATTTTCTGCGCGTCTCTCTGACACG 3117  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCCCTCTGCTACTCCATCTCTGAAGCAAGACGAGGGATGCTGGTGGGGCGCAAG 3177  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3178 GGCGCCCGCGGCTCTGCGCTCCGAGGGCGGTGACGTGGCTGTGCCACCAAGCATTTCTTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 CTCAAGCTGATCTGCACACCGTGTACCTAGTGCCTCTCTGGGGTCACTCAGGACAGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3298 CAGACGAGCTGAGTCCGAAGCTCCCGGGGACGACGCTGACTGCCCTTGAGGCCCGCAGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCCGCGACTGCGCTTCAGACTTCAAGACCATCTCTGGAC 3396

09963

ABX09963 standard; DNA; 3399 BP.

ABX09963;

23-JAN-2003 (first entry)

Human telomerase reverse transcriptase DNA fragment SEQ ID 28.

oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;  
prion; inhibition; human; ds.

**Homo sapiens.**

DE10100587-C1.

21-NOV-2002.

09--JAN-2001: 2001DE-1000587.

09-TAN-2001: 2001DE-1000587.

(RIBO-) RIBOPHARMA AG.

Kreutzer R. Limmer S. Rost S. Hadwiger P.

WPI: 2002-742209/81.

Inhibiting expression of target genes, e.g. oncogenes, in cells, by introduction of complementary double-stranded oligoribonucleotide, after treating the cell with interferon -

Disclosure: Page 30-31: 98pp; German.

This invention describes a novel method for inhibiting expression of a target gene by introducing into the cell that contains the target gene at least one oligonucleotide (dsRNAi) that has a double-stranded (ds) structure of not more than 49 consecutive nucleotides (nt), where at least a segment of one strand of the ds structure is complementary with the target gene and the cells are treated with interferon before introduction of dsRNAi. The method is used to inhibit expression of target genes, particularly oncogenes, cytokine genes, Id (not defined) protein genes, developmental or prion genes, or genes expressed in pathogenic organisms (particularly plasmidia) or in viruses or viroids (pathogenic in humans, animals or plants). Treating the cells with interferon greatly increases the extent to which dsRNA can inhibit expression of the target genes, and the effect is even greater when dsRNA are modified to increase their stability. ABX09936-ABX10075 represent gene fragments used to illustrate the method of the invention.

Sequence 3399 BP: 550 A: 1157 C: 1089 G: 603 T: 0 other;

ianment Scores:

ed. No.:	9.5e-234	Length:	3399
gment Scores:		Matches:	719
cre:	3505.00	Conservative:	132
cent Similarity:	72.94%	Mismatches:	260
st Local Similarity:	62.3%	Indels:	52
ery Match:	59.40%	Gaps:	13
	24		

-09-042-460-2 (1-1122) x ABX09963 (1-3399)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
1 ARGCCGGCGCTCCCGCTGCCAGCCGTCCTCTCCCTGCGCAGCCATACCCGAG 50  
21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
61 GTGCTCCCGCTGCCACGTCTGTCGCGCGCCTGGGCGCCCGCGGCTGGGTGTGCGAG 120  
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
21 CGCGGGGAGCCCGCGCGCTTTCGCGCGCTGTGTGCGACAGTCTGTGTGTGCGTGGCCGTGG 180

411 ValArgLeuLeuArgSerHisCysAsgPheArgThrAlaAsnGlnValThrAspAla 430  
::: :::  
1216 GGGGTGCTCTCTCAAGACGCACATGCCCGCTCGAGTGGG-----GTCACCCCACGA 1266  
431 Leu-----Asn 432  
1267 GCCGTGTCTGTGCCCGGAGAACCCCHAGGCTCTGTGGCGGCCCGAGGAGGAGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValYr 452  
1327 ACAGACCOCCTGCTCGCTGGTGACACTGCTCGGCAGACAAGCACGCGCCCTGGCAGGTGAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGCTTCGTGGCGGCTGCTCGCGCGGTGTGCCCCAGGCGCTCTGGGGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheLysAsnLeuLysPheIleSerLeuGlyLysTyrgLYls 492  
1447 AACGAACGCGCTCTCTCAGAACACCAAGAAGTTTCATCTCCCTGGGAGCATGCCAAG 1506  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CFTCTCGCTGCAGAGACTGACGTGGAAAGATGAGCGTGGGAGACTGCGCTTGGCTCGCAGG 1566  
513 SerProGlyLysAspAsgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCACCGGGTTGGCTGTCTTCGGCGCGCAGACACCGTCTCGTGAGAGATCTCGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrrValValGlnLeuLeuArgSerPhePheTyrr 552  
1627 AAGTTCCTCGCACTGGCTGATGAGTGTGTACGTGCTGCTGCTGCTCAGTCTTTCTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrrArgLysSerValTrpSer 572  
1687 GTCCAGAGACACAGTTTCAAAGAACAGCGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTGCAAGACATTGGAATCAGACAGCACTTGAAGAGGGTGACGTCGCGGAGCTGTGG 1806  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAGCAGAGGTTCAGCAGCATCGGAGACCGCGCCCTGCTGACGTCACAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrrSerMetGlyThr 632  
1867 TTCATCCCAGCGCTGACGGCTGCGCGATGTGTGAACATGACTAGCTCGTGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAAGCTTCCCGCAGAAAAGAGCGCGACGCTCACCTCGAGGGTGGAAGGCACTGCTTC 1986  
653 SerMetLeuAsnTyrrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 ACCGTGCTCACTACGAGGGCGCGCGCGCCCGCGCTCTCTGGGCGCTCTGTGCTGGGC 2046  
673 MetAsnAspIleTyrrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCCACAGGCGCTGGCGCACCTTCGCTCGTGTGCGGCGCCAGACCCG 2106  
693 ThrProArgMetTyrrPheValLysAlaAspValThrGlyValAspAlaIleProGln 712  
2107 CGCCTGAGCTGTACTTTGTCAAGGTGGATGTACGGGCGGTGACACCATCCCCCAG 2166  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrrCysile 732  
2167 GACAGGCTCAGGAGGTCAATGCCAGCATCATCAAA---CCCCAGAACACGTACTGCGGT 2223  
733 ArgGlnTyrrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 CGTCGGTATGCGGTGGTCCAGAGCGCGCCATGGGCGACGTCCGCAAGGCTTCAAGAGC 2283  
753 GlnValThrThrLeuSerAspLeuGlnProTyrrMetGlyGlnPheLeuLysHisLeuGln 772

Db	2284	CACGTCCTACCTTCAGACAGCCTCCACGCGGTACATGCGACAGTTCGTGGCTACCTCGAC	2343
Qy	773	AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn	792
Db	2344	GAGACC-----AGCCGCTGAGGGATGCGCTCGTCATCGACGAGAGCTCTCCCTGAAT	2397
Qy	793	GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValIys	812
Db	2398	GAGGCGACAGTCGCCCTCTTCGACGCTCTTCCTACGCTTCATGTGCGACACGCGCTGCGC	2457
Qy	813	IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr	832
Db	2458	ATCAGGGCAGTCTCTACGTCCAGTGCACGGGATCCCGCAGGGCTCCATCTCTCCACG	2517
Qy	833	LeuLeuCysSerLeuCysPheGlyAspMetGluAenLysLeuPheAlaGluValGlnArg	852
Db	2518	CTGCTCTGCGCCTGCTACTACGGCGCATCGAGAAACAAGCTGTTTTCGGGGATTCGCGCG	2577
Qy	853	AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln	878
Db	2578	GACGGGCTGCTCCCTGCGTGTGGTGAGATTTCTTGTGTGACACCTCACTCAACCAC	2637
Qy	873	AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn	892
Db	2638	CGCAAAACCTTCCTCAGACCCCTGGTCCGAGGTGTCCTCGATGCTGCGTGTGCTGAAC	2697
Qy	893	LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro	912
Db	2698	TTGCGGAAGACAGTGGTGAATCTCCCTGTAGAAGACGAGGCCCTGGGTGGCACCGCTTTT	2757
Qy	913	TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr	932
Db	2758	GTTCCAGATGCCGCCGCCACGCCCTATTCCCTCGTGGCGGCTGCTGCTGATACCCGGACC	2817
Qy	933	LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerIleuThr	952
Db	2818	CTGAGGTCGACAGCGCATCTCCAGCATATGCCGACCTCCATCAGACGCCAGTCTCACC	2877
Qy	953	PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg	972
Db	2878	TTCAACCGCGCTTCAGGCTGGAGGACATGCGTCGCAACTCTTTGGGGTCTTGGCG	2937
Qy	973	LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle	992
Db	2938	CTGAAGTGTACAGCTGTTCTCGATTTGCAAGTGAAACAGCCTCCAGACGGTGTGCACC	2997
Qy	993	AniIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu	1012
Db	2998	AACATCTACAAGATCCCTCGTGCAGCGGTACAGGTTTCAGCATGTGCTCGACGCTC	3057
Qy	1013	ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln	1032
Db	3058	CCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTCGCGGTCTCTGTACACG	3117
Qy	1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer	1052
Db	3118	GCTCCCTCTGCTACTCCATCCTGAAGACCAAGACGACGAGGATGCTCGTGGGGCCCAAG	3177
Qy	1053	GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu	1069
Db	3178	GCGCGCGCGGCCCTCTGCCCTCCGAGGCGGTGCAGTGGCTGTGCGCACCAAGCATTCCTG	3237
Qy	1070	LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla	1089
Db	3238	CTCAAGCTGACTCAGCACCGGTGCACCTACGTGCCACTCTCGGGGTCTCATCAGACAGCC	3297
Qy	1090	GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla	1109
Db	3298	CAGACGAGCTTGATCGAAGCTTCCCGGGAAGACGCTGACTGCTCGGCGCGCGACCC	3357
Qy	1110	AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp	1122

3358 AACCCGGCACTGCCTCAGACTTCAAGACCATCTCGGAC 3396

SULT 7

78144

ABV78144 standard; DNA; 3399 BP.

ABV78144;

15-NOV-2002 (first entry)

Human telomerase reverse transcriptase DNA SEQ ID NO 28.

RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic; virucide; protozoicide; gene; ds.

Homo sapiens.

WO200255693-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-EP00152.

09-JAN-2001; 2001DE-1000586.

28-OCT-2001; 2001DE-1055280.

29-NOV-2001; 2001DE-1058411.

07-DEC-2001; 2001DE-1060151.

(RIBO-) RIBOPHARMA AG.

Kreutzer R, Limmer S, Rost S, Hadwiger P;

WPI; 2002-590671/63.

Inhibiting expression of target gene, useful e.g. for inhibiting oncogenes, by administering double-stranded RNA complementary to the target and having an overhang

Claim 10; Page 129-130; 203pp; German.

The invention relates to inhibiting expression of a target gene (I) in a cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded structure of at most 49 consecutive bases. At least part of one strand (asi) of dsRNA1 is complementary to (I) and at least one end of dsRNA1 has an overhang of 1-4 nucleotides. The method is used to inhibit the expression of a wide range of genes, e.g. oncogenes, cytokine genes etc. in humans, also genes in Plasmodium or in viruses or viroids that are pathogenic for humans, animals or plants. Introducing an overhang into dsRNA greatly increases effectiveness for inhibiting gene expression, both in vivo and in vitro and also increases stability and thus the effective concentration inside the cell. The present sequence is that of a gene related to the invention.

Sequence 3399 BP; 550 A; 1157 C; 1089 G; 603 T; 0 other;

ignment Scores:

Seq. No.:	Length:	Score:
9.5e-234	3399	3505.00
Matches:	719	
Conservative:	122	
Mismatches:	260	
Indels:	52	
Gaps:	13	

09-042-460-2 (1-1122) x ABV78144 (1-3399)

Seq. No.:	Length:	Score:
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20	1122	3505.00
1 ATGCCGGCGCTCCCGCTCCGAGCGCTGCTCCCTGCTGCGCAGCCATACCGCGAG 60	1122	3505.00
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyValArgLeuValGln 40	1122	3505.00
61 GTGCTGCGCTGGCCACAGTTCGTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 120	1122	3505.00

Seq. No.:	Length:	Score:
41 ProGlyAspProTyrIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60	1122	3505.00
121 CGCGGGACCCGGCGGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCG 180	1122	3505.00
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuValGluLeu 80	1122	3505.00
181 GAGCAGCG 240	1122	3505.00
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100	1122	3505.00
241 GTGCGCGCGCTGCTGCGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300	1122	3505.00
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120	1122	3505.00
301 TTCGCGCTGCTGCGAGCG 360	1122	3505.00
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140	1122	3505.00
361 AGCTACCTGCCCAACACCGGTGACCGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420	1122	3505.00
141 LeuSerArgValGlyAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160	1122	3505.00
421 CTGCGCGCGCTGCG 480	1122	3505.00
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180	1122	3505.00
481 CTG 540	1122	3505.00
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200	1122	3505.00
541 GCCACTCAGCG 594	1122	3505.00
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220	1122	3505.00
595 -----TGGCAACGGCGCTGGAACCATACCTACCTACCTACCTACCTACCTACCTACCT 633	1122	3505.00
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240	1122	3505.00
634 GTCCCTCTGCGCGCTGCG 693	1122	3505.00
241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259	1122	3505.00
694 CTGCGCTTGGCCCAAGAGCG 753	1122	3505.00
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272	1122	3505.00
754 GGGCAGGGTCTGCG 813	1122	3505.00
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292	1122	3505.00
814 GTGCTGCTACCTGCCAGA-----CCCGCGCAAGAGCCACCTCTTTGGAG 858	1122	3505.00
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311	1122	3505.00
859 GGTGCGCTCTCTGCGCAGCG 918	1122	3505.00
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330	1122	3505.00
919 CCCCACATCCACATCCG 978	1122	3505.00
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350	1122	3505.00
979 GAGACCAAGCACTTCTCTACTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1035	1122	3505.00
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370	1122	3505.00
1036 CTACTCAGCTCTCTGAGCG 1095	1122	3505.00
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390	1122	3505.00
1096 CTGGGTTCAGGCG 1155	1122	3505.00
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410	1122	3505.00

|||||TACTGGCAATGCGGCCCTCTTTCTGGAGTGTCTGGGAACACGCGCAGTGCCTTAC 1215  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAlaGlnGlnValThrAspAla 430  
1216 GGGGTGCTCTCAAGACCACTGCGCGTGGAGCTGG-----GTCAACCCAGCA 1266  
431 Leu-----Asn 432  
1267 GCGGTGTCTGTGCGCGGGAAGCCCGAGGGCTCTGTGCGGCGCCCGAGGAGGAC 1326  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValThr 452  
1327 ACAGACCCCGCTGCGCTGGTGCAGCTCTCGGCAGACAGACCGCCCTGGCAGGTGTAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGTTCGTGCGGGCTGCTGCGCGCGCTGGTGCAGGCTCTGGGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheLeuAsnLeuLysPheLeuSerLeuGlyLysTrpGlyLys 492  
1447 AACGAACCGCTTCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1506  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCCTGCTGCGAGGCTGACGTGGAAGATGAGCGTGGGACTGCGCTTGGCGCGCAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCAGGGGTGGTGTGTCGCGCGCAGAGCACCGTCTGCGTGGAGAGATCTCGGC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTrpValValGlnLeuLeuArgSerPheThr 552  
1627 AAGTTCCTGCACTGCTGATGAGTGTGTACGTGCTGCGTGTGAGTGTCTTCCTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTrpArgLysSerValTrpSer 572  
1687 GTACGCGAGACCGCTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGCGAGTGGCGAGCTGTGC 1806  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 512  
1807 GAAGCAGAGGTGAGCAGCATCGGGAAGCCAGCCCGCTCTGTGAGCTCAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTrpSerMetGlyThr 632  
1867 TTCATCCCAAGCTGACGGGCTGGGCGGATGTGAACATGGACTACGTCTGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAACGTTCCGAGAGAAAGAGGCGGAGGCTCACTCCCTCGAGGGTGAAGGCACCTGTT 1986  
653 SerMetLeuAsnTrpGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 ACGTGTCTCACTACAGCGGGCGGCGCGCCCGCTCTCTGGCGCTCTGTGTGGG 2046  
673 MetAsnAspIleTrpArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCCACAGGGCTGGCGCACCTTCTGTGTGTGTGGGGCCAGGACCCG 2106  
693 ThrProArgMetTrpPheValLysAlaAspValThrGlyAlaTrpAspAlaIleProGln 712  
2107 CGCCTGAGCTGTACTTGTCAAGTGGATGTACGGGCGCGTACGACACCATCCCGCCAG 2166  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTrpCysIle 732  
2167 GACAGGCTCAGGAGTTCATCGCGAGCATCATCAA---CCCCAGAACAGTACTGGTG 2223  
733 ArgGlnTrpAlaValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752

Db 2224 CGTCGTATCCGCTGTGTCCAGAGGCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAGC 2283  
Qy 753 GlnValThrThrLeuSerAspLeuGlnProTrpMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2284 CAGCTCTCTACTTGCAGACCTCCAGCGGTATCATGCGACAGTTCGTGGCTCCTGCAG 2343  
Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2344 GAGACC-----AGCCCGCTGAGGGATGCGTCTCATCGACGAGCTCTCTCCCTGAAT 2397  
Qy 793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
Db 2398 GAGCCAGCAGTGGCTCTCTCGACGCTCTTCACGCTTATGTGCCACCCAGCCGCTGGC 2457  
Qy 813 IleGlyAspArgCysTrpThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
Db 2458 ATCAGGGGCAAGTCTCTACGTCCAGTCCAGTCCAGGGATCCCGCAGGGCTCCATCTCTCCAG 2517  
Qy 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
Db 2518 CTGCTCTGCAGCCTGTGTCTACGCGCAGCATGAGAACAGAGCTGTTTGGGGGATTCGGCG 2577  
Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
Db 2578 GACGGCTGCTCTCTGCTTGTGGTGAATGTTCTTGTGTGACACCTCACCCTCACCAC 2637  
Qy 873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTrpGlyCysMetIleAsn 892  
Db 2638 GCGAAACCTTCTCAGGACCTCTGGTCCGAGGTGTCCTGAGTATGGCTGCTGGTGAAC 2697  
Qy 893 LeuGlnLysThrValAlaAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
Db 2698 TTGCGAAGACAGTGTGTAACTTCTCTGTAGAGACGAGGCCCTGGGTGGCACGGCTTTT 2757  
Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2758 GTTCAGATGCGGCGCCACCGCCTATTTCCTCTGGTGGGCTGCTGCTGGATACCCGACC 2817  
Qy 933 LeuGluValPheCysAspTrpSerGlyTrpAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2818 CTGGAGGTGCGAGCGCACTACTCAGCTATGCCGCGACCTCCATCAGACGCCAGTCTCAC 2877  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
Db 2878 TTCAACCGCGCTTCAAGGCTGGGAGAACATGCTGCGCAAACTCTTTGGGCTCTTGG 2937  
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2938 CTGAAGTGTCAAGCTGTTTCTGGATTTCAGGTGAACAGCCTCCAGACGGTGTGCACC 2997  
Qy 993 AsnIleTrpLysIlePheLeuLeuGlnAlaTrpArgPheHisAlaCysValIleGlnLeu 1012  
Db 2998 AACATCTACAGATCTCTCTGCTGCGAGCGTACAGGTTCACGCGATGTGTGTGCGCTC 3057  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3058 CCATTTCATCAGCAGTTTGGAGAACCCCACTTTTCTCTGGCGGTCACTCTTGACAGC 3117  
Qy 1033 AlaserCysCysTrpAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3118 GCCTCTCTGTACTCCATCTCTGAAGCCAGAACGAGGATGTGCTGGGGGCGCAAG 3177  
Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTrpGlnAlaPheLeu 1069  
Db 3178 GCGCCGCGCGCTCTGCGCTCCGAGGCGGTGAGTGTGCTGTGCCAACCAAGCATTCCTG 3237  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTrpLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3238 CTCAAGTGTACTCGACACCGTGTCACTACGTGCCACTCTCTGGGGTCACTCAGGACGCC 3297  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3298 CAGACGCGAGTGTGGAAGCTCCCGGGGACGACGCTGACTGCTGCTGAGGCGCGACGCC 3357





1096 CTGGGTCCAGGCCCTGGATGCCAGGAGCTCCCGCAGAGTTGCCCGCCCTGCCCCAGCGC 1155  
 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 1156 TACTGGCAAAATGGCGCCCTGTTCTTGGAGCTGCTTGGGAACCAACGCGAGTGGCCCTAC 1215  
 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 1216 GGGGTGCTCTCAAGACGCACTGCCCGCTCGAGCTGCG-----GTACCCACGCA 1266  
 431 Leu-----Aan 432  
 1267 GCCGGTGTCTGTGCCCGGAGAACCCCGAGGCTCTGTGTGGCGGCCCGCAGGAGGAGAC 1326  
 433 ThrSerProHisLeuMetAspLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 1327 ACAGACCCCGTCCGCTGGTGGAGCTCTCCGCGACAGACAGCCCTGGAGAGTGTAC 1386  
 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 1387 GCGTTCGTGGGGCTGCTGCTGGCGCGCTGCTGCCCGCCCTCTGGGCTCCAGGCAC 1446  
 473 AsnGluArgArgPhePheLysAsnLeuLysPheLysLeuGlyLysTyrGlyLys 492  
 1447 AACGAACCGCGCTTCTCCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1506  
 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 1507 CTCTCGCTGAGGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGCTGGTGGCGAGG 1566  
 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLysLeuAla 532  
 1567 AGCCACAGGGTGGCTGTGTCCGGCGCAGACACCGCTGCGGTGAGGAGATCTCTGGCC 1626  
 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 1627 AAGTTCCTGCACTGGCTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1686  
 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
 1687 GTCACGGAGACACGTTTCAAAAGACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1746  
 573 LysLeuGlnSerIleGlyValArgGlnHisGluArgValArgValArgLeuArgGlnLeuSer 592  
 1747 AAGTGTCAAAAGCATTGAATCAGACACACATTTGAAGAGGTTGACGTGGCGGAGCTGTGC 1806  
 593 GlnGluValValArgHisHisGlnAspThrTrpLeuAlaMetProLysArgLeuArg 612  
 1807 GAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCCCGCTGTGTGAGCTCCAGACTCCGC 1866  
 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
 1867 TTCATCCCAAGCCTGACGGGCTCGCGCCGATGTGAACATGGACTACGTCTGTGTGGAGCC 1926  
 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
 1927 AGACCTTCCCGAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGCATTGTTTC 1986  
 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
 1987 AGCGTGTCAACTACGAGCGGGCGCGCGCGCCCGCGCTCTCTGGCGCCCTGTGTGTGGGC 2046  
 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
 2047 CTGGACGATATCCACAGGGCGCTGGCGACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2106  
 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
 2107 CCGCTGAGCTGTACTTGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2166  
 713 GlyLysLeuValGluValAlaAsnMetLysArgHisSerGluSerThrTyrCysIle 732  
 2167 GACAGGCTCACGGAGGTTCATCGCGAGCATCATCAAA-----CCCCAGAACACGTACTCGGTG 2223

Qy 733 ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
 Db 2224 CGTGGTATGCGGTGGTCCAGAGGCCCGCCATGGGCACGTCGCGAAGCGCTTCAAGAGC 2283  
 Qy 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
 Db 2284 CAGGTCTCTACCTTGACAGACCTCCAGCGGTATACGACAGTTCGTGGCTCACCTGCAG 2343  
 Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
 Db 2344 GAGACC-----AGCCCGCTGAGGATCCGTCGTCATCGACGAGAGCTCTCTCCGTGAAT 2397  
 Qy 793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
 Db 2398 GAGCCAGCAGTGGCTCTTTCGACGTCTTCTACCGTCAATGTGCCACACCGCGTGGC 2457  
 Qy 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
 Db 2458 ATCAGGGGCAAGTCTCTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCTCTCCACG 2517  
 Qy 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
 Db 2518 CTGCTCTGAGCCCTGTGCTACCGCGACATGAGAGAACAGCTGTTTGGCGGGATTCGGCGG 2577  
 Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
 Db 2578 GACGGGCTGCTCTCGCTTGTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCAC 2637  
 Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
 Db 2638 GCGAAACCTTCTCTCAGGACCTGTGCTACCGCGACATGAGAGAACAGCTGTTTGGCGGGATTCGGCGG 2697  
 Qy 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
 Db 2698 TTGGGAGACAGTGTGTGAACCTTCCCTGTGTGGAGAGCGGCCCTGGTGGCAGCGCTTTT 2757  
 Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
 Db 2758 GTTCAGATGCGCGCCACCGCCCTATTCCCTGTGTGGCGCTGCTGTGTGGATACCGCGACC 2817  
 Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
 Db 2818 CTGGAGTGTGAGAGCAGCTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACC 2877  
 Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
 Db 2878 TTCAACCGCGGCTTCAAGGCTGGAGAACATGCTGCGCAAACTCTTTGGGCTCTTGGCG 2937  
 Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
 Db 2938 CTGAAGTGTACAGCTGTGTTCTGGATTTCAGGTGAACAGCTCCAGCGGTGTGACCC 2997  
 Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
 Db 2998 AACATCTACAAGATCTCTCTGCTGCAGCGGTACAGGTTTCAGCATGTGTGTGTCAGCTC 3057  
 Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
 Db 3058 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTCTGCGCGTCTCTCTGACACG 3117  
 Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
 Db 3118 GCTCTCTCTGTCTACTCTCATCTCTGAAGCAAGACGAGGAGATGCTCTGGGGGCCAAG 3177  
 Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
 Db 3178 GGGCGCGCGGCCCTCTGCGCCCTCCGAGCGGTGACGTGGCTGTGCGCACCAAGCATTCCTG 3237  
 Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 Db 3238 CTCAAGTGTACTGACACCGCTGTCACTACGTGGCCACTCTCTGGGTCACTCAGGACAGCC 3297

1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
 3298 CAGACGAGCTGAGTCGGAAGCTCCCGGACGAGCGTACTGCTGGAGCGCGAGCC 3357  
 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 3358 AACCCGGCACTGCCCTCAGACTTCAGACCATCTGGAC 3396

MULT 9  
 727876

AAV27876 standard; cDNA; 3798 BP.

AAV27876;

25-MAR-2003 (updated)

12-OCT-1998 (first entry)

Human telomerase protein 2 (TP2) full-length cDNA.

TP2; human; telomerase protein 2; cancer; AIDS; ageing; therapy;  
 SS.

Homo sapiens.

Key Location/Qualifiers  
 CDS 2..3466  
 /tag= a

WO9821343-A1.

22-MAY-1998.

13-NOV-1997; 97WO-US21248.

15-NOV-1996; 96US-0751189.

11-JUN-1997; 97US-0873039.

16-OCT-1997; 97US-0951733.

(AMGE-) AMGEN CANADA INC.

(AMGE-) AMGEN INC.

Harrington LA, Robinson MO;

WPI; 1998-297946/26.

P-PSDB; AAV27876.

New nucleic acid encoding human telomerase protein-2 - used for  
 regulating telomerase activity, e.g. for treating cancer or acquired  
 immune deficiency syndrome

Claim 1c; Fig 9; 150pp; English.

This full-length cDNA clone codes for a human telomerase protein  
 2 (TP2, see AAW61350), a novel protein of the telomerase complex.  
 The sequence was deduced from overlapping partial clones #32 (see  
 AAV27876) and TP2-15 (see AAV27875), which were obtained from a human  
 colon tumour cell line LIM1863 cDNA. Expressing TP2 genes in a  
 cell is used to increase telomerase activity and thus proliferation  
 for treatment of e.g. HIV infection, AIDS and ageing disorders,  
 while expressing an inactive mutant of TP2 (or molecule antisense  
 to the gene) is used to decrease telomerase activity, e.g. for  
 treatment of cancer. TP2 polypeptides can also be used to screen  
 for agents that inhibit TP2 activity or its binding to TRP1 (see  
 AAW61347) or telomerase RNA, potentially useful therapeutically,  
 also to raise specific antibodies useful in immunoassays and  
 therapeutically as inhibitors. Nucleic acid fragments are used  
 as diagnostic probes for detecting/quantifying TP2 DNA. Also  
 contemplated are transgenic animals in which the TP2 gene has been  
 inactivated or is overexpressed. TP2 polypeptides are administered  
 i.v., s.c. or orally, or they are delivered from engineered cells  
 or gene therapy vectors.  
 (Updated on 25-MAR-2003 to correct PR field.)

SQ Sequence 3798 BP; 613 A; 1311 C; 1212 G; 662 T; 0 other;

Alignment Scores: 1.11e-233 Length: 3798  
 Pred. No.: 3505.00 Matches: 719  
 Score: 3505.00 Conservative: 122  
 Percent Similarity: 72.94% Mismatches: 260  
 Best Local Similarity: 62.36% Indels: 52  
 Query Match: 59.40% Gaps: 13  
 DB:

US-09-042-460-2 (1-1122) x AAV27876 (1-3798)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 Db 68 ATGCGGCGCTCCCGCTGCGAGCGGTGGCTCCCTGCTGGCAGCCACTACCGCAG 127  
 QY 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
 Db 128 GTGCTGCGCTGGCCACGTTGCTGCGCGCTGGGGCCCGGCGGTGGCGGTGGCGAG 187  
 QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethIstTp 60  
 Db 188 CGCGGGACCGCGCGCTTCCCGCGCTGTGGCCAGTGGCTGGTGGCGCTGGCTGG 247  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 Db 248 GAGCACGGCGCGCGCGCGCGCGCTCCCTCCCGCAGGTGTCTCTGCTGAAGGAGTGG 307  
 QY 81 ValAlaArgValValGlnArgLeuGlyGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 308 GTGCGCGGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGGCGGAGAACGTGTGGCTTCG 367  
 QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerSerValArg 120  
 Db 368 TTCGCGCTGTGGAGCGCGCGCGCGCGCGCGCTTACACCGCTTACACCGCGTGGCG 427  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaTrpMetLeu 140  
 Db 428 AGCTACCTGCCCAACACGCTGACCGAGCTGCGGGGAGCGGGGCGGTGGGGGCTGCTG 487  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 488 CTGCGCGCGTGGCGGACGAGCTGTGTTCACCTGTGGCAGCTGGCGCTCTTTGTG 547  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 548 CTGTGGCTCCAGCTCGCTTACAGGTGTGGCGCGCGCGCTGTACCGCTCGCGCT 607  
 QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 608 GCCACTCAGCGCGCGCGCGCGCGCTAGT---GGACCCCGAGGCGGTCTGGGA--- 661  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 Db 662 -----TGCAACGGCGCTGGAACCATAGCTCAGGAGGCGCGG 700  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240  
 Db 701 GTCCCTTCGGCGCTGCGAGCGCGCGGTGCGAGAGCGCGCGGCGGAGTCCCGCAAGT 760  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 761 CTGCGTGTCCCAAGAGCGCGCGCGGTGGCGCTGCCCTGAGCGGAGCGGAGCGCGCTT 820  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTp--- 272  
 Db 821 GGGCAGGCGTCTCGGCGCGCGCGCGGAGCGCGGTGGACCGAGTGGCTGTTCTGT 880  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 881 GTGGTGTCACTGTGCAGA-----CCCGCGGAGAGGACCACTCTCTTGGAG 925  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311

[illegible]

1994 AGAAGCGTTTCGACGAGAAAAAGAGGGCCGAGCGTCTCACTCGAGGGGTGAAGCAGCTGTC 2056

653 SerMetLeuAsnTyrGIuArgThrLysHisProHisLeuMetGlySerValIeuGly 672  
:::  
2054 ACGGTGCTCAACTACGAGCGGGCGCGCGCGCGCGCTCTCGGGCGCTCTGTGCTGGC 2113

673 MetAsnAspIleTyrArgGthrTrpArgAlaPheValLeuArgValArgAlalaLeuaspGln 692  
:::  
2114 CTGGACGATATCCACAGGGGCTCGGCACCTTGCTGCTGCTGCTGCTGGGCCCGACGACCG 2173

693 ThrProArgMetTyrPheValLysAlaAspValThrGIyValATyrAspAlaIleProGln 712  
:::  
2174 CCGCTGAGCTGTAATTGTCAAGTGGATGTACGGGCGGTACGACCATCCCCAC 2233

713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
:::  
2234 GACAGGCTCACGGAGGTCATCGGCACGATCATCAA--CCCAGAACAACGTACTGGGT 2290

733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
:::  
2291 CGTGGTATGCGTGGTTCAGAGGCCGCCCATGGCAGCTCCGACAGGCCTTCAAGAGC 2350

753 GlnValThrThrLeuSerAppLeuGlnProTyrMetGlyGlnPheLeuLysHisIeuGln 772  
:::  
2351 CACGCTCTACTCTGACAGACCTCCAGCGGTACATGCCAGAGTTCGTGGCTCACCTGCAG 2410

773 AppSerAppAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleserMetAsn 792  
:::  
2411 GAGACC-----AGCCCGCTGAGGATGCGCTGCTCATCGACAGAGTCTCTCCCTGAA 2454

793 GluSerSerSerLeuPheAppPhePheLeuHisPheLeuArgHisSerValVallys 812  
:::  
2465 GAGGCGACGAGTGGCTCTTCAGAGCTTCCTCCTACGCTTCATGTGCACACGCGCTGGC 2524

813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerIeuSerThr 832  
:::  
2525 ATCAGGGCAAGTCTCAGTCCAGTGCAGGGATCCCGAGGGCTCATCTCTCTCCACG 2584

833 LeuLeuCysSerLeuCysPheGlyVasPmetGluAsnLysLeuPheAlaGluValGlnArg 852  
:::  
2585 CTGCTCTGACGCTGTGTACGGCGCATGAGAGAACAGCTGTTGCGGGATTTCGGCGG 2644

853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisIeuLeuaspGln 872  
2645 GACGGCTGCTCTCGCTTGGTGGATGATTTCTGTGTGTGCACACCTCACCTCACCCAC 2704

873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2705 GCGAAAACCTTCTCTCAGGACCTGGTCCGAGGTGTCCCTGAGTAGTGTGGTGGTGAC 2764

893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
:::  
2765 TTGCGAGACAGAGTGGTGAACCTTCCCTGTATAAGAGAGAGCGGCCCTGGGTGGACGCGTTTT 2824

913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLysLeuLeuAspThrGlnThr 932  
2825 GTTCAGATCGCGGCCACGGCTATTCCCTGTGTGGGCTGTCTGTGTGATACCCGGACC 2884

933 LeuGluValPheCysAspTyrSerGIyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2885 CTGGAGGTGCAGAGGACTACTCCAGCTATGCCCGGACCTCCCATCAGAGCCAGTCTCAC 2944

953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValIeuArg 972  
2945 TTCACCGCGGCTTCACGGCTGGAGGAAATGCGTCGCAAACTCTTTGGGTCTTCGGG 3004

973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
3005 CTGAAGTGTACAGGCTTTTCTGATTTGAGGTGAAACAGCTCCAGAGGTGTGCACC 3064

993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnIleu 1012  
3065 AACATCTACAGATCTCTGCTCAGCGCTACAGTTTCAGCATGTGTGTGCGAGTCT 3124



596 GCCACTCAGCGCCGCCGCCACACGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGCAACGGCGCTCGAACCACATACGCTCAGGAGGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCCCTGGCGCTGCCAGCCCGCGGTGCGAGGAGCGCGGGCGCATGCGCCGAGT 748  
241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluPro--- 259  
749 CTGCGGTGCCACAGAGCGCCAGCGGTGCGCTGCCCTGAGCGGAGCGACGCCGCT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGGTCTCGGGCCACCCGGGAGAGCGGTGACCGAGTACCGTGTTCGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTCACTTGCAGCA-----CCCGCGAAGACCACTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTTGGCTCTCTGGCACCGCGCCACTCCACCCATCCGTGGGCCCGCCAGCACCGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCATCGGGCCACACGCTCCCTGGGACAGCGCTTCTCCCGGTAGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTCTCTACTCTCTCAGCGCAC---AAGGAGCAGCTGCGCGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCCGAGCTGAGTGGCGCTCGAGGCTCTGGAGACCACTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTCCAGGCTTGGATGTCAGGAGCTCCCGCAGGTTGCGCCGCTGCCCGCAGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGCAATCGCGCCCTGTCTTCTGAGTGTCTTGGAAACCAACCGCGAGTGCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTCTCTCAAGACGCACTGCCCGTGGAGTGCG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCGTGTCTGTCCCGGAGAACCCAGGCTCTGTGGCGGCCCGGAGGAGGAC 1381  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACNAGCCCGCTGCTGTGTGAGTGTCTTCCGACGACACAGCCCTTGCAGGTTGAT 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTTCGTGCGGGCTGCTGCGCGCGGTGTGTCGCCCGAGGCTCTGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AAGAAACCGCTCTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCAGCCAA 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTGCTCAGAGTGTACGTGGAGATGAGCGTGGCGGACTGCGCTTGGCTGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluAlaIleLeuAla 532  
1622 AGCCAGGGGTGGTGTGTCTCCGCGCCGACAGACACCGTCTGCTGCTGAGAGATCTCGCC 1681

533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGCTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGGACCAAGTTCAAAAGAACAGGCTCTTTTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGlnLeuSer 592  
1802 AAGTTCGAAAGCATTTGAATCAGACGACACTTGAAGAGGCTGAGCTGCGGAGCTGTGC 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAGCAGAGGTACGAGCATCGGAAGCAGCGCCCTGCTGCTGACGTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTGACCGGCTGCGCGCATTTGTGAACATGACACTACGCTGCGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGACAGAAAAGAGCGCGCGCTCCTCCTCGAGGGGTGAAGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGCGCGCGCGCTCTCTGCGCGCTCTGTGTGGGC 2101  
673 MetAsnApIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGCGCTGCGCACCTCTGCTGCTGCTGCGGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGGCTGAGTGTACTTGTCAAGGTGATGTGCGGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGCGAGGTCTATCGCCAGCATCATCAAA---CCCCAGAACACGACTGCGTG 2278  
733 ArgGlnTrpAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCAGAGCGCCCGCATGGGACGTGCGCAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACTTTCACAGACCTCCAGCGGTACATGCGACAGTTCGTGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGGATGCGCTGTCATGAGCAGAGCTCTCCCTGGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGCTGCGCTTTCGACGCTCTTCTACGCTTCTATGTCACACCGCGCTGCGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTCTACGTCCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACGCTGCTGACGGGACATGAGAACAGAGCTGTTTGGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 CACGGCTGCTCTCTGCGTGTGTGATGATTTCTTGTGTGACAGCTCACCTCACCCAC 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 CGAAAAACCTTCTCTCAGACCCCTGCTCCGAGGTGTCTCTGATGCTGCTGCGTGGTGAAC 2752



893 LeuGlnLysThrValAlaSerPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
 |||||  
 2753 TTGCGAAGACAGTGTGAACCTTCCCTGTAGAACAGAGCCCTGGGTGCACGGCTTTT 2812  
 |||||  
 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
 |||||  
 2813 GTTCAGATCGCGGCCACCGGCTATTCCCTGTGTGGCTGTCTGTGGATACCGGACC 2872  
 |||||  
 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerLeuLysThrSerLeuThr 952  
 |||||  
 2873 CTGGAGGTGCAGAGGAGTACTCTCAGCTATATCCCGGACCTCCATCAGAGCAGCTTCACC 2932  
 |||||  
 953 PheGlnSerValPheLeuAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
 |||||  
 2933 TTCACCGCGGCTTCAGGCTGGAGGACATGTGTGCACACTCTTTGGGTCTTGGGG 2992  
 |||||  
 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
 |||||  
 2993 CTGAAGTGTACAGGCTGTCTTCGATTTGCAGGTGAACAGCTCCAGACGGTGTGCACC 3052  
 |||||  
 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
 |||||  
 3053 AACATCTCAAGATCTCTCTGTGAGCGGTACAGTTTCAGCATGTGTGTGTCAGCTC 3112  
 |||||  
 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
 |||||  
 3113 CCATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGTGCGGTCACTCTGACACG 3172  
 |||||  
 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
 |||||  
 3173 GCTCCCTCTGTCTACTCCATCTCAAGCCAGAACGAGGAGTGTCTGGGGGCCAAG 3232  
 |||||  
 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
 |||||  
 3233 GGGCGCGCGGCTCTGCTCCCGAGCGCGGAGTGTGTGTCACCAAGCATTCCTG 3292  
 |||||  
 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 |||||  
 3293 CTCAGCTGACTCGACACCGGTCACTACGTGCCACTCTCTGGGTCACTCAGACAGCC 3352  
 |||||  
 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
 |||||  
 3353 CAGACGAGCTGAGTCGAGAGCTCCCGGGGACGAGCTGACTGCCCTGGAGCGCGAGCC 3412  
 |||||  
 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 |||||  
 3413 AACCGCGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3451  
 |||||

SULT 11  
 Z30154

AAZ30154 standard; cDNA; 4015 BP.

AAZ30154;

26-JAN-2000 (first entry)

cDNA encoding a human telomerase reverse transcriptase (TRT).

Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;  
 dendritic cell; telomerase activity; cancer cell; proliferating cell;  
 immunological destruction; telomerase; cancer; proliferation disease;  
 ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 56..3454

/\*tag= a

/product= "telomerase reverse transcriptase"

W09950392-A1.

07-OCT-1999.

XX 30-MAR-1999; 99WO-US06898.  
 PF |||||  
 XX 31-MAR-1998; 98US-0112006.  
 PR |||||  
 XX (GERO-) GERON CORP.  
 PA |||||  
 XX Gaeta FCA;  
 PI |||||  
 XX WPI; 1999-610845/52.  
 DR P-PSDB; AAY43621.  
 XX Eliciting an in vivo immune response for prevention and treatment of  
 cancers -  
 PT |||||  
 XX Disclosure; Fig 2; 26pp; English.  
 PS |||||  
 XX The present sequence encodes a human telomerase reverse transcriptase  
 CC (TRT) polypeptide. The protein is used in the method of the invention.  
 CC The specification describes a method for activating a T lymphocyte,  
 CC comprising contacting the T lymphocyte with a dendritic cell that  
 CC expresses a TRT peptide in the context of a MHC class I or MHC class  
 CC II molecule. The protein causes induction of an in vivo immunological  
 CC response to telomerase activity. Cancer cells are characterized by  
 CC expression of endogenous TRT gene and the presence of detectable  
 CC telomerase activity. Therefore, by eliciting a specific immune response  
 CC to TRT or to TRT-expressing cells, it is possible to selectively target  
 CC proliferating cells for immunological destruction. The method is used  
 CC for eliciting an in vivo immune response to telomerase by activating  
 CC a T lymphocyte, and is useful for prevention and treatment of cancers and  
 CC other proliferation diseases/conditions.

XX SQ Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.2e-233 Length: 4015  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: Gaps: 13

US-09-042-460-2 (1-1122) x AAZ30154 (1-4015)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 |||||  
 Db 56 ATGCGCGCGCTCCCGCTGCGCGCGCTGCTGCTGCTGCGCGCGCTGCTGCTGCGCGAG 115  
 |||||  
 QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
 |||||  
 Db 116 GTGCTGCGCTGGCCACGTTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGAG 175  
 |||||  
 QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 |||||  
 Db 176 CGCGGGGACCGCGCGCTTCCCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 235  
 |||||  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 |||||  
 Db 236 GAGCAGCG 295  
 |||||  
 QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 |||||  
 Db 296 GTGGCGCGAGTGTCTGAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
 |||||  
 QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerSerValArg 120  
 |||||  
 Db 356 TTCGGCTGTGTGAGCGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
 |||||  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 |||||  
 Db 416 AGCTACCTGCCCAACACGTCACGACGCACTGGCGGGGAGCGGGCGGTGGGGCTGCTG 475  
 |||||  
 QY 141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 |||||



	CDS	56..3454 /*tag= a
FT	XX	WO9950386-A2.
PN	XX	07-OCT-1999.
PX	XX	31-MAR-1999; 99WO-US07097.
PF	XX	31-MAR-1998; 98US-0052864.
PR	XX	03-AUG-1998; 98US-0128354.
PP	XX	(GERO-) GERON CORP.
PA	XX	Morin GB;
PI	XX	WFI; 1999-610842/52.
DR	XX	P-PSDB; AAY32090.
DT	XX	New catalytic polypeptide and polynucleotide, useful for increasing catalytic activity in a cell
PT	XX	Disclosure; Fig 2; 24pp; English.
PS	XX	This is the nucleotide sequence of cDNA encoding human telomerase reverse transcriptase (hTERT, see AAY32090). Human telomerase is a target for diagnosing and treating diseases relating to cell proliferation and senescence, such as cancer, or for increasing the proliferative capacity of a cell. A claimed method for increasing the proliferative capacity of a vertebrate cell, involves introducing especially a human or other mammalian cell, involves introducing into the cell a recombinant hTERT polynucleotide encoding an hTERT variant in which residues 192-323, 200-323, 192-271, 200-271, CC 222-240, 415-450, 192-323 and 415-450, or 192-271 and 415-450 are deleted. A claimed method for reducing telomerase activity in a cell involves introducing a recombinant polynucleotide encoding an hTERT variant having a deletion of amino acids 192-450, 560-565, CC 637-660, 638-660, 748-764 or 1055-1071. The polynucleotides are obtained by mutagenesis of the hTERT coding sequence.
SQ	XX	Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;
	Alignment Scores:	
	Pred. No.:	Length: 4015
	Score:	Matches: 719
	Percent Similarity:	Conservative: 122
	Best Local Similarity:	Mismatches: 260
	Query Match:	Indels: 52
	DB:	Gaps: 13
	US-09-042-460-2 (1-1122) x AAZ20279 (1-4015)	
QY	1	MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db	56	ATGCCGGCGCTCCCGCTGCCAGCGGTGCCTGCTCGCGAGCACCTACCGCGAG 115
QY	21	ValTrpProLeuAlaThrPheValArgLeuGlyProGluArgArgValGln 40
Db	116	GTGTCGCCGTGGCCACGTTCGTGCGCGCCCTGGGGCCCCAGGGCTGGCGCTGGTGAG 175
QY	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60
Db	176	CGCGGGGACC CGGGCTTTCCGCGCGCTGGTGGCCAGCTGCTGGTGGCTGGCTGG 235
QY	61	GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
Db	236	GACGACG 295
QY	81	ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
Db	296	GTGGCCCCAGTGTCTGACAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 355
QY	101	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120

356 TTCCGCTGTGTGACAGGGGCGCGGGGCGCCCGGAGGCGCTTTCACCAACGCGTGC 415  
121 serTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrrpMetLeuLeu 140  
416 AGTACTACTGCCAACACGGGTGACCCAGCAGCATGCGGGGAGCGGGGGTGGGGGCTGCTG 475  
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGCGCGCGTGGGCGACAGACGTGCTGGTTCACTTCGTGGCAGCTCGCGCGCTTTTGTG 535  
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGTGTGCTCCAGCTCGCCTACCAAGGTGTGCGGCCCGCGCTGTACCAAGCTCGCGCGT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCGCGGCCCGCCGACACGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
650 -----TCCGAACCGGCGCTCGAACATACGCTCAGGAGGCGGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
689 GTCCCGCTGGCGCTGCAGCCCGGTGCGAGGAGCGCGGGGCGAGTGCACGCCGAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
749 CTGCGGTGTGCCAAGAGCGCCAGCGCTGCGCTGCGCTCCTAGACGCGAGCGAGCGCGGT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGTCTTGGGCCACCGGCGCAGGAGCGGTGACCGAGTACCGTGGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGAACCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLeuHisLysProSer 311  
914 GTGGGTCTCTGGCAGCGGCCACTCCACCCACTCGTGGGGCGGCCAGCACACGCGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCATCCACATCGCGGCCACCACTGCTGGGACACGCTTGTCCCGGTGTACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgLysAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTCTACTCTCAGCGCAG---AAGGAGCAGCTCGCGCCCTCCTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGTCTCTGAGGCCACGCTGACTGGCGCTCGAGGCTGTGGACCACTCTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTCTCCAGGCCCTCGATCGCAGGGAATCCCGCAGGTGTCGCCCGCTCCCGCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnIleValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATATCGGCCCTCTGTTCGTGGAGCTGCTGGGAACACGCGCAGTGCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCCTCAAGACGCACTGCGCGCTGCGAGCTGCG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCGTGTCTGTGCCGGAGAACCCCGAGGCTCTGTGGCGGCCCGGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452

1382	ACAGACCCCGCTCGCTGGTGCAGCTGCTCCGCGACACAGCAGCAGCCCTCGCAGCGTGTAC	1444
453	GlyPheLeuAArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis	472
1442	GCCTTCGTGCGGGCTTCGCTCGCGCGCTGGTGCCTCCCGCAGGCTCTCGGGGCTCCAGGCAC	1501
473	AsnGluAArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyValYsrGlyLys	492
1502	AACGAACCCGCTTCCTCAGGAACAACAAGAGTTCACTCCCTCGGGAAGCATGCCAAG	1561
493	LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer	512
1562	CTCTCGCTGCAGAGCTGACGTGGAGATGACGCTGCGGACTGCGCTGGCTGCAGCAGG	1621
513	SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleuAla	532
1622	AGCCACGAGGGTGGCTGTGTTCGCGCGCAGCAGCACCGTCTCGTGAGAGATCCTGGCC	1681
533	ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr	555
1682	AGTTCTCTGCTGCTGT	1741
553	IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer	572
1742	GTCCGAGAGACCAAGTTCTAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC	1801
573	LysLeuGlnSerIleGlyValAArgGlnHisLeuGluAArgValAArgLeuArgGluLeuSer	592
1802	AGTTGTCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGTCACTCGCGAGCTGTGCG	1861
593	GlnGluGluValAArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg	612
1862	GAAGCAGAGGTTCAGCGAGCATCGGAAGCCAGCGCCCTGCTGACGTCACAGCTCCGCG	1921
613	PheIleProLysProAsnGlyLeuAArgProIleValAsnMetSerTyrSerMetClyThr	632
1922	TTCACTCCCAAGCTGACGGGCTCGCGCCGATTTGTAACATGGACTACGTGTGGGAGCC	1981
633	ArgAlaLeuGlyAArgArgLysGlnAlaGlnHisPheThrGlnAArgLeuLysThrLeuPhe	652
1982	AGAACGTTCCGACAGAAAGAGGCGCGAGCGCTCACTCGAGGGGTGAAGCAGCTGTC	2041
653	SerMetLeuAsnTyrGluAArgThrLysHisProHisLeuMetGlySerSerValLeuGly	672
2042	AGCGTCTCACTACAGCGGCGCGCGCGCGCGCTCTCTCGGGCGCTCTGTGTGGGCG	2101
673	MetAsnAspIleTyrAArgThrTrpAArgAlaPheValLeuAArgValAArgAlaLeuAspGln	692
2102	CTGACAGATATCCACAGGCGCTGCGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGT	2161
693	ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln	712
2162	CGCGCTGAGCTGACTTGTCAAGTGGATGTACGCGGCGGTAGACACCATCCCCCAG	2221
713	GlyLysLeuValGluValAlaAlaAsnMetIleAArgHisSerGluSerThrTyrCystile	732
2222	GACAGCTCACGAGGTTCATCGCAGCATCATCAA--CCCGCAACACGACTGTGGTG	2278
733	ArgGlnTyrAlaValValAArgAspSerGlnGlyValHisLysSerPheArgArg	752
2279	CGTGGTATCGGTGTCACGAAGCGCCCATGGGACGTCGCGAAGGCTTCAAGAGC	2338
753	GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln	772
2339	CAGCTCTCTACCTTGACAGACCTCCAGCCCTACATCGCAGTGTGTGGCTCACCTGAG	2398
773	AspSerAspAlaSerAlaLeuAArgAsnSerValValIleGluGlnSerIleSerMetAsn	792
2399	GAGACC-----AGCCCCCTGAGGGATGCGCTGTGTATCGACGACAGCTCTCTCCCTGAAT	2452
793	GluSerSerSerLeuPheAspPheLeuHisPheLeuAArgHisSerValValLys	812
2453	GAGCCAGCAGTGGGCTCTTCGAGCTTCTCAAGTCTTCTCAAGTTCATGTGCGCACACGCGCTGGC	2512

```

813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 932
|||
2513 ATCAGGGGCAAGTCTCAGTCCAGTCCAGGGATCCCGAGGGCTCCATCCTCTCCACG 2572
|||
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 952
|||
2573 CTGGCTCGAGCTGTCTACGGCGACATGGAGAACAGCTGTTTGGGGGATTCGGCGG 2632
|||
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 972
|||
2633 GACGGGCTCTCTCGTCTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCAC 2692
|||
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
|||
2693 GCGAAACCTTCTCAGGACCTGTGCGAGGTGTCCCTGAGTATGCTGGTGGTGAC 2752
|||
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
|||
2753 TTGGGAGACAGTGTGAATCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTTT 2812
|||
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuAspThrGlnThr 932
|||
2813 GTTCAGATCCCGGCCACCGCCTATTCCCTGGTGGGCCCTGCTGCTGATGATCCCGAAC 2872
|||
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
|||
2873 CTGGAGTGCAGAGCGACTACTCCAGCTATGCCGACCTCCATCAGAGCCAGTCTCACC 2932
|||
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
|||
2933 TTCAACCGCGGCTTCAAGGCTGGGAGGACATGCGTGCACAACTCTTTGGGGTCTTCGGG 2992
|||
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
|||
2993 CTGAAGTGTACAGCTGTCTTCTGATTTGAGGTGACACGCTCCAGACGGTGTGCACC 3052
|||
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
|||
3053 AACATCTACAAGATCTCTCTGTGTCAGGCGTACAGGTTTCACGATGTGTGCTGCAGCTC 3112
|||
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
|||
3113 CCAATTCATCAGCAAGTTTGGAGAACCCACATTTTCTTGGCGGTCACTCTGCACAG 3172
|||
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
|||
3173 GCCTCCCTCTGTCTACTCTCTCTGAAAGCAAGAACGAGGAGTGTCTGCTGGGGGCCAAG 3232
|||
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
|||
3233 GGGCGCGCGGCGCTCTGCCCTCCGAGGCGGTGCAGTGGCTGTGGCCACAGCATTCCTG 3292
|||
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
|||
3293 CTCAAGCTGACTGCACACCGTGTCACTACGTGCCACTCTCTGGGGTCACTCAGACAGCC 3352
|||
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
|||
3353 CACAGCAGCTGAGTGGAGCTTCCCGGGAGCAGCGCTGACTGCTGGAGGCGCGAGCC 3412
|||
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
|||
3413 AACCGGCACTGCCCTCAGACTTCAGACCACTCCTGGAC 3451
|||

```

RESULT 13  
AAZ00724

AAZ00724 standard; DNA; 4015 BP.

AAZ00724;

06-OCT-1999 (first entry)

# Human telomerase catalytic domain DNA.

Telomerase; catalytic domain; human; quantitation; tumour cell; melanoma;  
body fluid; metastases; T-cell lymphoblastoma; chronic myeloid leukemia;  
acute lymphatic leukemia; melanoma; pulmonary carcinoma; colon cancer;  
breast cancer; ss.

Homo sapiens.

DE19804372-A1.

05-AUG-1999.

04-FEB-1998; 98DE-1004372.

04-FEB-1998; 98DE-1004372.

(DAHM/) DAHM M W.

Dahm MW;

WPI; 1999-431408/37.

Quantifying tumor cells by amplifying mRNA encoding the catalytic  
subunit of telomerase

Example; Fig 1A-B; 26pp; German.

This invention describes a novel method for the quantitation of tumour  
cells in a body fluid which comprises (1) enrichment or isolation of  
tumour cells in the sample, (2) amplification of mRNA from these cells  
that encodes the catalytic subunit of telomerase and (3) quantifying  
the amount of amplified mRNA. The method is applied to tumour cells  
derived from (micro)metastases, e.g. associated with a wide range of  
tumours such as T-cell lymphoblastoma, chronic myeloid or acute  
lymphatic leukemia, melanoma, pulmonary carcinoma, cancer of colon or  
breast etc. This sequence encodes a human telomerase protein catalytic  
domain.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Alignment Scores:	1.2e-233	Length:	4015
Pred. No.:	3505.00	Matches:	719
Score:	72.94%	Conservative:	122
Percent Similarity:	62.36%	Mismatches:	260
Best Local Similarity:	59.40%	Indels:	52
Query Match:	20	Gaps:	13

US-09-042-460-2 (1-1122) x AAZ00724 (1-4015)

QY	1	MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu	20
Db	56	ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGGCAGCCACTACCGCGAG	115
QY	21	ValTrpProLeuAlaThrPheValArgLeuGlyProGluArgLeuValGln	40
Db	116	GTGCTCGCTGGCCACGCTTCTGCGCGCGCTGGGGCCCGAGGCTGGCGCTGGTGCGAG	175
QY	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp	60
Db	176	CGCGGGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGGCTGGTGGCGCTGGCCTGG	235
QY	61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
Db	236	GACGACG	295
QY	81	ValAlaArgValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
Db	296	GTGCGCGCGAGTGTGCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	355
QY	101	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120

356 TTCGGCGTCTGGACGGGGCCCGCGGGGGCCCCCGGAGGCGCTTCCACACACAGCGTGGC 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
141 AGTACTCTCCCAACACGGGTGACCGACACTGCGGGGGAGCGGGGGCTGGGGCTCTG 475  
141 LeuSerArgValGlyAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGGCGCGGTGGCGACACGCTGTGTGTACCTGTCTGGACGCTGGCGCTTTGTG 535  
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla 180  
536 CTGGTGGCTCCAGCTGCCCTACCAAGTGTGGGGCGCGCGCTGTACCACTGGCGCT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCACCTCAGCG 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
650 -----TGCGAAGCGGCGCTGGAACCATAGCGTCAGGGAGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCTCTGGCGCTGCCACCGCGGTGGAGAGCGCGCGGGCGAGTGCACCGCGAAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259  
749 CTGGCGTTCACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
809 GGGCAGGGGTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSer 292  
869 GTGGTGTACCTGCCAGA-----CCGCGCGAGAGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGGCGCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCACATCGCGGCGCACCGCTCCCTGGGACACGCGCTGTGCCCGGTGACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCAGCTTCTCTACTCTCAGCGGAC---AAGGAGCAGCTGCGCGCTCTCTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCAGCGCTGACTGGCGCTCGGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGGTTCAGGCCCTTGATGCGAGGACTCCCGCAGGTGGCGCGCGCTGCCCGCGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCGCGCTGTCTTGTGGAGTGTCTGGGAAACCAAGCGCGAGTGCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGCACTGCCGCTGCGAGTGGC-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTGCGCGGAGAACCCAGGCGCTCTGTGGCGCGCGCGCGCGAGGAG 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGGCTGTGAGCTGTCTGCGCGACGACGACGACGCGCGCGCGGTGATC 1441

QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTTCGGTGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501  
QY 473 AsnGluArgArgPhePheLeuLysAsnLeuLysPheLeuLysTyrGlyLys 492  
Db 1502 AACGAACCGCTTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTGCGAGAGCTGACGTGGAAGATGAGGTGCGGAGCTGCGCTGGCTGGCGAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCACAGGGGTGGCTGTGTTCGCGCGCGAGACACCGTCTGCTGAGGAGATCCTGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
Db 1682 AAGTTCTCTGCACTGGCTGATGAGTGTGTAGCTGTCTGAGCTGTCTCTTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTACGGAGACCACTTTCATAAAGACAGGCTCTTTTCTACCGAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGAGCTGTGC 1861  
QY 593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTGAGCGAGCATCGGAAGCGCGCGCGCGCGCGCGCTGTGCTGAGCTCCAGACTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCATCCCAAGCTGACGGCTGCGCGCGCTGTTGTGAACATGAGCTAGCTGCGGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAAGTTTCCGACAGAAAAAGAGCGCGCGCTCTACCTCGAGGGTGAAGGCACTGTTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCTCACTAGAGCGCGCGCGCGCGCGCGCGCTGCTGGCGCGCTGTGCTGGGC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGACGATATCCACAGCGCTGCGGACCTTCTGCTGCTGCGGGGCCAGGACCG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGTACTTGTCAAGTGTGTGACGGCGCGGTACGACACCATCCCCCAG 2221  
QY 713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGCTCAGGAGTCACTCCCGCAGCATCAAAA---CCCCAAGAACAGTACTCGGTG 2278  
QY 733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
Db 2279 CGTGGTATGCGGTGTCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2338  
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2339 CAGCTCTTACTTTCAGAGACTCCAGCGGTACATGGACAGTTCGTGGCTCACTCGAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2399 GAGACC-----AGCCCGCTGAGGATCGCGTCTCATCGAGCAGAGCTCCTCCCTGAAT 2452  
QY 793 GluSerSerSerLeuPhePhePheLeuHisPheLeuArgHisSerValValLys 812  
Db 2453 GAGCCAGAGTGGCTCTTCGAGCTTCTCTAGCTTCTATGTGCCACACCGCGCGTGC 2512



```
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
|||
2513 ATCAGGGGCAAGTCTTACATCCAGTCCAGGCGATCCCGAGGGCTCCATCTCTCCACG 2572
|||
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAenLysLeuPheAlaGluValGlnArg 852
|||
2573 CTGCTCTGCAGCCTGTGCTACGGCGACATGAGAACAGCTGTTTGGCGGATTCGGCGG 2632
|||
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872
|||
2633 GACGGGCTGCTCTCGGCTTGGTGGATGATTTCTTGTGGTGCACCTCCTCACCCAC 2692
|||
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
|||
2693 GCGAACAACCTTCTCAGGACCTGTGCTCCAGAGTGTCCCTGAGTATGGTGGTGGTGAAC 2752
|||
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
|||
2753 TTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTT 2812
|||
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
|||
2813 GTTCAGATGCGGCGCCACGCGCTATTCCCTGCTGGTGGGCTCTCTCTGATACCGGAC 2872
|||
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
|||
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATGCCGAGCCTCCATCAGAGCCAGTCTCAC 2932
|||
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
|||
2933 TTCAACCGCGGCTTCAAGGCTGGGAGAACATGCTGTCGGAACCTTTTGGGGTCTTGGCG 2992
|||
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
|||
2993 CTGAAGGTGCAGACGCTGTCTTCTGGATTTCAGGTCAACAGCCTCCAGACGGTGTGCACC 3052
|||
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
|||
3053 AACATCTACAAGATCTCCCTGCTGCGAGCGGTACAGGTTTCAGCGATGTGTCTGAGCTC 3112
|||
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
|||
3113 CCATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGGCTCATCTCTGACAG 3172
|||
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
|||
3173 GCTTCCCTCTGCTACTCCATCTCTGAAGCCAGAACGCGAGGATGTGCTGGGGCCCAAG 3232
|||
1053 GlySer-----PheProGluAlaAlaHisIleLeuLysTyrGlnAlaPheLeu 1069
|||
3233 GCGCGCGCGCGCTCTGCTCTCGAGCGCGTGCAGTGGCTGTGCACCAAGCATTCCTG 3292
|||
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
|||
3293 CTCAGCTGACTCGACACCGGTGTCCCTAGTGCCTCCTCGGGTCACTCAGGACAGCC 3352
|||
1090 GlnLysLeuLeuCysArgGlyLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
|||
3353 CAGACGAGCTGAGTGGAGAGCTCCCGGGAGCAGCGTGTACTCTCCCTGGAGCGCGAGCC 3412
|||
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
|||
3413 AACCGGCACTGCTCAGACTTCAGACCATCTCGGAC 3451
|||
RESULT 14
AAH45901
D AAH45901 standard; DNA; 4015 BP.
C
X AAH45901;
C
T 06-SEP-2001 (first entry)
X
X Human hTERT gene.
E
```

XX Human; hTERT; telomerase; catalytic subunit; mRNA quantitation; detection; beta-region; diagnosis; cancer; ds.

XX Homo sapiens.

Key	Location/Qualifiers
exon	1..274
FT	/tag= a
FT	/number= 1
FT	275..1628
FT	/tag= b
FT	/number= 2
FT	1629..1924
FT	/tag= c
FT	/number= 3
FT	1825..2005
FT	/tag= d
FT	/number= 4
FT	2006..2185
FT	/tag= e
FT	/number= 5
FT	2186..2341
FT	/tag= f
FT	/number= 6
FT	complement (2309..2325)
FT	/tag= g
FT	/note= "primer SYC1076 (AAH45902) binding site"
FT	complement (2311..2325)
FT	/tag= h
FT	/note= "primer SYC1118 (AAH45905) binding site"
FT	2342..2437
FT	/tag= i
FT	/number= 7
FT	complement (2345..2374)
FT	/tag= j
FT	/bound_moiety= "hybridisation probe CS3"
FT	/note= "AAH45908"
FT	complement (2427..2456)
FT	/tag= k
FT	/bound_moiety= "hybridisation probe CS12"
FT	/note= "AAH45906"
FT	2438..2523
FT	/tag= l
FT	/number= 8
FT	complement (2458..2487)
FT	/tag= m
FT	/bound_moiety= "hybridisation probe CS1"
FT	/note= "AAH45907"
FT	2489..2506
FT	/tag= n
FT	/note= "primer SYC1097 (AAH45904) binding site"
FT	2524..2637
FT	/tag= o
FT	/number= 9
FT	2615..2631
FT	/tag= p
FT	/note= "primer SYC1078 (AAH45903) binding site"
FT	2638..2709
FT	/tag= q
FT	/number= 10
FT	2710..2898
FT	/tag= r
FT	/number= 11
FT	2899..3025
FT	/tag= s
FT	/number= 12
FT	3026..3087
FT	/tag= t
FT	/number= 13
FT	3088..3212
FT	/tag= u
FT	/number= 14

exon 3213..3350  
 /tag= v  
 /number= 15  
 exon 3351..4015  
 /tag= w  
 /number= 16  
 BP1108789-A2.  
 20-JUN-2001.  
 15-DEC-2000; 2000EP-0127228.  
 16-DEC-1999; 99US-0465491.  
 (HOFF ) HOFFMANN LA ROCHE & CO AG P.  
 Chang SP, Santini CD;  
 WPI; 2001-376930/40.  
 Quantitating expression of mRNA encoding hTERT, the catalytic subunit of telomerase, as an indicator of cancer, by amplifying RNA using primers complementary to hTERT gene sequence and quantitating amplified products  
 Claim 1; Page 5-7; 29pp; English.  
 The present sequence is that of the hTERT gene encoding the catalytic subunit of the human telomerase, comprising 16 exons, which is useful in a method for quantitating hTERT mRNA. The method is useful for detecting the presence of beta-region (a 182 nucleotide region consisting of exons 7 and 8) of the hTERT-mRNA in a human sample for diagnosis and prognosis of cancer. The method provides an accurate measure of telomerase activity by selectively measuring mRNA that encodes an active hTERT protein.  
 Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;  
 Alignment Scores:  
 ad. No.: 1.2e-233 Length: 4015  
 ore: 3505.00 Matches: 719  
 cent Similarity: 72.94% Conservative: 122  
 st Local Similarity: 62.36% Mismatches: 260  
 ary Match: 59.40% Indels: 52  
 Gaps: 13  
 -09-042-460-2 (1-1122) x AAH45901 (1-4015)  
 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCGCGCGCTCCCGCTGCGAGCGCGTGCCTCCCTGCTGCGCAGCCACTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgGluValGln 40  
 116 GTGCTCCGCTGGCCACGTTGCTGCGCGCTGCGGCGCCCGAGGCTGCGGCTGGTCAG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGCGGGAGCCCGCGCTTCGCGCGCTGTTGGCCAGTGCCTGGTGTGGTCCCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GAGCGAGCGCGCGCCCGCCCTCTTCCGCGAGGTGCTGCTGCAAGAGAGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGCGCCGAGTGTGCGAGAGGTGTGCGAGCGCGCGGCGAGACGTGCTGCGCTTCGGC 355  
 101 PheGluLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTGCGCTGTGTGACG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140

Db AGCTACTGCCCCAACCGTGCACGACACTGCGGGGAGCGGGCGGTGGGGCTGCTG 475  
 Qy LeuSerArgValGlyAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db CTGCGCGCGTGGGCGACGACGTGCTGGTTCCTGCTGCGACGCTGCGCGCTTTTGG 535  
 Qy LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db CTGGTGGCTCCAGCTGCGCTTACAGGTGTGGGGCGCGCTGTACAGCTGGCGCT 595  
 Qy ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db GCCACTCAGCG 649  
 Qy AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220  
 Db TCGGAACGGGCTTGAACCATAGCTCAGGAGGCGCGG 688  
 Qy LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
 Db GTCCCGCTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
 Qy ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259  
 Db CTGCGGTGGCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 808  
 Qy HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
 Db GGGCAGGGGTCTCTGGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
 Qy ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
 Db GTGTGTGTCACCTGCGCAGA-----CCGCGCGAAGAACCCACCTCTTTGGAG 913  
 Qy GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db GGTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
 Qy SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db CCCCCATCCATCCG 1033  
 Qy GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db GAGACCAAGCACTTCTCTACTCTCTCAGGGGAC---AAGAGAGAGCTGCGCGCTCTTC 1090  
 Qy LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 Db CTACTCAGCTCTGTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1150  
 Qy LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
 Db CTGGTTCCAGCGCTCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210  
 Qy TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 Db TACTGGCAATGCT 1270  
 Qy ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db GGGGTGCTCTCAAGACGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1321  
 Qy Leu-----Asn 432  
 Db GCCGTGTCTGTGCGCGGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1381  
 Qy ThrSerProProHisLeuMetAspLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db ACAGACCG 1441  
 Qy GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472

1442 GGCTTGTGGCGGCGCTGCCTTCGCGCGCGCTGGTGCCGCCCAAGGCGCTCTGGGGGCTCAGGGCAC 1501

473 AsnGluArgArgPhePheLysAsnLeuLysPheIleSerLeuGlyLysTyroGlyLys 492  
:::  
1502 AACGAACGCGCTTCCTCAGGAACACCAAGAAGTTTCATCTCCTGGGAAGCATGCCAAG 1561

493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCCTCGTGCAGGAGCTGCAGTGAAGATGAGCGTGGGACTGCGCTTGGCTGGCGAGG 1621

513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGGTGGCTGTGTCCGGCCGAGACACCGCTCGCTGAGGAGATCTCTGGCC 1681

533 ThrPheLeuPheTrpLeuMetAspThrTyrvValValGlnLeuLeuArgSerPhePheTyR 552  
1682 AAGTTCTCTGCATCTGGCTGATGAGTGTCTAGCTCGTGCAGTGTCTCAGGCTTTCTTTTAT 1741

553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyRArgLysSerValTrpSer 572  
1742 GTCACGGAGACCAGTTTCAAAGAACAGAGCTCTTTTTCTACCGGAAGAGTGTCTGGAGC 1801

573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGGAATCAGACAGCATTTGAAGAGGTGCAGCTGGGGAGCTGTGC 1861

593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTTCAGGACGATCGGGAAAGCAGGCCCGCCCCCTGTGTGACGTCCAGACTCCC 1921

613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSertYrSerMetGlyThr 632  
1922 TTTCATCCCAAGCCTGACGGCTCGCGCCGATTGTGAACGACTACGTCTGCGGGAGCC 1981

633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCCGCAGAGAAAGAGCGCGAGCGTCTCACCTCGAGGGTGAAGGACACTGTT 2041

653 SerMetLeuAsnTyRGlUArgThRLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCACTACAGCGGGCGGGCGCCCGCCGCTCTCTGGGGCGCTCTGTGCTGGGC 2101

673 MetAsnAspIleTyRArgThRTripArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGAGCATATCCACAGGCGCTGCGCACCTTCGTGTGCTGTGGCGGCCCAAGACCCG 2161

693 ThrProArgMetTyRPheValLysAlaAspValThrGlyAlaTyRaspAlaIleProGln 712  
2162 CGCGCTTGACCTGTACTTGTCAAGTGTGTGTACGGCGCGGTACGACACCATCCCGCAG 2221

713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyRCysIle 732  
2222 GACAGCTCACGGAGTTCATCGCCAGCATCATCAAA---CCCCAGAACAGTACTGCGTG 2278

733 ArgGlnTyRAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGTGTATCGCGTGTCCAGAACCGCCCATGGCACGTCGCCAAGGCGCTTCAAGAGC 2338

753 GlnValThrLeuSerAspLeuGlnProTyRMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTCTACCTTACAGACCTTCAGCGGTACATGCGACAGTTCTGTGGCTCACCTGCAG 2398

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGAGATCCGTCGTATCGAGCAGAGGTCTCTCCCTGAAT 2452

793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGAGTGGCCCTCTTCAGAGCTTCTTCCTACGCTTCATGTGCCACACCGCGCTGCC 2512

813 IleGlyAspArgCysTyRThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGGAAGTTCATCTCAGTGCAGTGCAGGGGATCCCGCAGGGCTTCATCTCTCCAGC 2572

QY	833	LeuLeuCysSerLeuCysPheGlyAspMetGluAenLysLeuPheAlaGluValGlnArg	852
DB	2573	CTGCTCTGCAGCTGCTGCTACGCGCAGCATGGAGAAACAAGCTGTTTGCGGGATTTGCGCGG	2632
QY	853	AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln	872
DB	2633	GACGGCTGCTCCTCGGTTTGGTGGATGATTTCTTTGGTGACACTCACTCACCAC	2692
QY	873	AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn	892
DB	2693	CGCAAAACCTTCTCTCAGACCCCTGGTCCAGAGTGCCCTGAGTAGGCTGCGTGGTGAC	2752
QY	893	LeuGlnLysThrValValAenPheProValGluProGlyThrLeuGlyGlyAlaAlaPro	912
DB	2753	TTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTTT	2812
QY	913	TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr	932
DB	2813	GTTTCAGATGCGCGGCCACAGGCCATTATTCCTCGTGGCGCTGCTGCTGGATACC	2872
QY	933	LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr	952
DB	2873	CTGGAGGTGCAGACGGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGCTCACC	2932
QY	953	PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg	972
DB	2933	TTCAACCCGCGCTTCAAGGCTGGGAGGAACATCGCTGCGCAAACTCTTTGGGGTCTTCGGG	2992
QY	973	LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle	992
DB	2993	CTGAAGTGTCAACGCTGTTCTTGGAATTTGCAGTGAACAGGCTCCAGACGGTGTGCACC	3052
QY	993	AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu	1012
DB	3053	AACATCTCAAGAATCCTCTGCTGCAGCGGTACAGGCTTTCAAGCATGTGTCTGCAGCTC	3112
QY	1013	ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln	1032
DB	3113	CCATTTCATCAGCAAGTTTGGAGAACCACCAATTTTCTCGCGGTGATCTCTTGACAGC	3172
QY	1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer	1052
DB	3173	GCTCCTCCTGCTACTCCATCCTGAAGCCAGAACCCAGAGGATGCTGCGGGGCCCAAG	3232
QY	1053	GlySer-----PheProProGluAlaAlaHisIleTyrLeuCysTyrGlnAlaPheLeu	1069
DB	3233	GCGCGCGCGCCCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG	3292
QY	1070	LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla	1089
DB	3293	CTCAAGCTGACTCGACACCGGTGCACCTACGTGCCACTCCTGGGTGTCATCAGGACAGCC	3352
QY	1090	GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla	1109
DB	3353	CAGACGACGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGCCCGCAGCC	3412
QY	1110	AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp	1122
DB	3413	AACCCGCGACTCCCTCAGACTTTCAGACCATCCTGGAC	3451

RESULT 15  
AAAD46821  
ID AAAD46821 standard; cDNA; 4015 BP.

RESULT 15

AAD46821  
 ID AAD46821 standard; cDNA; 4015 BP.

AC AAD46821;

DT 27-JAN-2003 (first entry)

XX  
DE Human telomerase reverse transcriptase (TERT) cDNA.

Carbohydrate antigen; alpha(1,3)galactosyltransferase; alpha1,3CT; TERT;  
KW  
XX  
transgenic: alpha(1,2)fucosyltransferase; alpha1,2FT; human; enzyme;  
KW  
XX

356 TTGCGGCTGCTGGACGGGGGCCCGGGGGCCCCCGAGGCTTCACCAACGAGGTGGC 415

453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472  
1442 GGCCTTCGTGGCGGCTTCCTGCGCGGCTGGTGGCCCCCAGGCGCTTCGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuTyrLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AACGACCGCGCTTCCTCAGGAACACCAAGAGTTCACTCCCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGTGCAGAGCTGAGCTGGAAGATGAGCGTCGCGGACTGGCGCTTGGCTGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGGTGGCTGTGTTCGCGCGCGCAGAGCACCGCTCTGCGTGAGAGATCCTGGCC 1681  
533 ThrPheLeuPheTyrPheLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACCTGGCTGATGAGTGTGTACGTCTGTCGAGCTCTCAGGTCTTCTTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572  
1742 GTACGGAGACCAAGCTTTCAAAGAACAGAGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTGTGCG 1861  
593 GlnGluValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgGluArg 612  
1862 GAACAGAGTTCAGGAGCATCGGAGAGCCAGGCGCGCCCTCTGACGTCCAGATCCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCTGACGGCTGCGCGCGATTGTGAACATGACTACGTCTGTCGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGCGAGAGAAAGAGCGGCGGCGCGCTCTCTGCGGCGCTCTGTGTGGGC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACGAGCGGCGCGCGCGCGCTCTCTGCGGCGCTCTGTGTGGGC 2101  
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCGCTGCGCACCTTCGTCTGCTGCTGCGGCGCGAGACCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGTGTGATGTACGGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 CACAGGCTCAGGAGGTCAATCGGCGAGCATCAAA---CCCCAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGATGCGGTGGTCCAGAGCGCGCCATGGGACGTCCGAGGCGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACTTTCAGACACCTCCAGCCGTATCATGCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetHis 792  
2399 GAGACCC---AGCCCGCTGAGGGAGTCCGTCTCATCGAGCAGAGCTCCCTCCCTGAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCCAGCTGGCCCTTCGACGTCTTCTACGCTTCTCATGTGCCACCCAGCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832

2513 ATCAGGCGCAAGTCTCTAGCTCCAGTGCAGGGGATCCGCGAGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGlnAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTCAGACCTGTGTACGGCGCATGAGAACAGCTGTTTGGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTCGGTTGGTGGATGATTTCTTGTGTGACACCTCACCCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTCTCTCAGGACCTGCTCGAGGTCCTGAGTATGGCTCGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTTCGGAGACAGTGGTGAATTTCTCTGAGACAGGAGCCCTGCTGCTGGTGGACGGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCCGGGCGCCACGGCTATTCCCTGCTGGTGGCGCTGCTGCTGGATACCCGAGC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATCCCGGACCTCCCATCAGAGCCAGTCTCAC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGGCTTCAAGGCTGGGAGGAACATGCTGCGAACTCTTTGGGCTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCGCTGTTCTGGATTGAGGTGAACAGCCCTCCAGAGCGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCTGAGCGGTACAGGTTTCACGCATGTGTGTGTCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTTCATCAGCAAGTTTGAAGAACCCCATTTTCTCGCGCGTCATCTCTGCACACG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCTCTGTACTTCATCTCTGAAAGCCAGAACCGCAGGATGTGCTGGGGGCCAAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisTyrPheCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGGCGCTCTGCGCTCCGAGGCGGTGAGTGGCTGTGCGCCACCAAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTCAGTCGACCGGTGTACCTACGTGCCATCTCTGGGGTCACTCAGGACGAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGCGCTGAGTGGAGGCTCCCGGGGACGAGCTGACTGCTGCGCTGGAGCGCGAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGCGCATCCCTCAGACTTCAAGACCATCTCTGGAC 3451

Search completed: February 1, 2004, 15:20:47

Job time : 880 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

on: February 1, 2004, 15:05:59 ; Search time 144 Seconds  
(without alignments)  
3439.110 Million cell updates/sec

le: US-09-042-460-2

fect score: 5901

quence: 1 MTRAPRCAPVRLSRVRE.....TILKAADPALSTDPQTILD 1122

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

urched: 569978 seqs, 220691566 residues

al number of hits satisfying chosen parameters: 1139956

imum DB seq length: 0

cimum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

mmand line parameters:

>DEL=frame.p2n.model -DEV=xlh

>/cgn2\_1/USPTO.spool/US09042460/runat 01022004 150552 17040/app.query.fasta\_1.1287

>Issued\_Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

>OPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi

>LST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

>DE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

>SER=US09042460 @CGN 1 1.71 @runat 01022004 150552 17040 -NCPU=6 -ICPU=3

>WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

>V TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

>APEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

abase :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles.seq:\*

#### SUMMARIES

ult	No.	Score	Query	Match	Length	DB	ID	Description
1	3505	59.4	3396	3	US-08-974-549A-638			Sequence 638, App
2	3505	59.4	3396	3	US-08-974-549A-639			Sequence 639, App
3	3505	59.4	3396	3	US-08-974-549A-640			Sequence 640, App
4	3505	59.4	3396	3	US-08-974-549A-641			Sequence 641, App
5	3505	59.4	3396	3	US-08-974-549A-642			Sequence 642, App
6	3505	59.4	3451	3	US-08-974-549A-721			Sequence 721, App
7	3505	59.4	4015	3	US-08-851-843A-224			Sequence 224, App
8	3505	59.4	4015	3	US-08-974-549A-1			Sequence 1, Appli
9	3505	59.4	4015	3	US-08-854-050-224			Sequence 224, App
10	3505	59.4	4015	4	US-09-430-323-224			Sequence 224, App
11	3505	59.4	4015	4	US-09-572-423B-3			Sequence 3, Appli
12	3505	59.4	4015	4	US-09-128-354-1			Sequence 1, Appli

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 343, App  
Sequence 173, App  
Sequence 292, App  
Sequence 173, App  
Sequence 173, App  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 100, App  
Sequence 266, App  
Sequence 100, App  
Sequence 100, App  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 30, Appli  
Sequence 8, Appli  
Sequence 62, Appli  
Sequence 62, Appli  
Sequence 8, Appli  
Sequence 68, Appli  
Sequence 111, App  
Sequence 68, Appli  
GENERAL INFORMA  
Sequence 111, App  
Sequence 1, Appli  
Sequence 109, App

#### ALIGNMENTS

#### RESULT 1

US-08-974-549A-638  
; Sequence 638, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997



PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/911,312  
 FILING DATE: 14-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,951  
 FILING DATE: 14-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/915,503  
 FILING DATE: 14-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US97/17618  
 FILING DATE: 01-OCT-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US97/17885  
 FILING DATE: 01-OCT-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph Ted  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002610US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 638:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3396 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 FEATURE:  
 NAME/KEY: 1..3396  
 LOCATION: 1..3396  
 OTHER INFORMATION: /note= "HTRT-encoding sequence employing  
 OTHER INFORMATION: alternative codon distributions for  
 OTHER INFORMATION: E. coli (all Genes)"  
 I-08-974-549A-638  
 Alignment Scores:  
 Seq. No.: 0 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Percent Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 Gaps: 13  
 I-09-042-460-2 (1-1122) x US-08-974-549A-638 (1-3396)  
 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 1 ATGCCGGCGCGCGCGCTGCGCGCGGTGCGCAGCCTGTCGCGAGCCATTATCGCGAA 60  
 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgValGln 40  
 61 GTGTCGCGCGCGCGCTTGTGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCG 120  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 121 CGCGCGCATCGCGCGCGTTTCGCGCGCTGTCGCGCAGTCGCTGCTGTCGCGCGTGG 180  
 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 181 GATCG 240  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100

Db 241 GTGGCGCGCGCTGTCGAGCGCTGTGGAACGCGCGCGGAAAAACGTCGTGGCGTTTGGC 300  
 QY 101 PheGluLeuLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 Db 301 TTTGCGCTGCTGGATGCG 360  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 Db 361 AGCTATCTCTGCCGAACACCGTGACGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 421 CTGCGCGCGCTGGCGGATGATGCTGTCGTCATCTGTCGCGCGCGCGCGCGCGCGCG 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 481 CTGCTGGCGCGCGCTGCGCGTATCAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCG 540  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 541 GCGACCCAGCG 594  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 Db 595 -----TGGGAACGCGCGCTGGAACCATAGCTGCGCGCGCGCGCGCGCGCGCGCG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrIleArgHisLeuSerLeuThrSerThrSer 240  
 Db 634 GTGCGCGCTGGCGCTGCG 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 694 CTGCGCGCTGCGGAACG 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 754 GCGCAGGCGAGCTGGCGCGCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 814 GTGTGTGAGCCCGCGCGCG-----CCGCGGGAAGAAGCGACGAGCGCTGGAA 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 859 GCGCGCTGAGCGCGCACCGCGCATCCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 919 CGCGCAGCACCG 978  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 979 GAAACCAACATTTTCTGTATAGCAGCGCGCAT---AAAGAACAGCTGCGCGCGAGTTT 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 Db 1036 CTGCTGAGCAGCTGCG 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 Db 1096 CTGGCAGCG 1155  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1156 TATTTGACATGCGCGCGCTGTTCTGGAACCTGCTGGGCAACCATGCGCGCGCGCGTAT 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 Db 1216 GCGGCTCTGCTGAACCACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
 QY 431 Leu-----Asn 432





[illegible]

1867 TTATATCCGAAAAACCGAGCGCTGTGGCTCGATGCTGTTAACTAGCATGCACTAGTGTGGTGCT 1321

633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652

1927 CGTACTTCCTCGTGAATAAACCGTGTGTAACGCTGTGACCTCCCGGTAAAGCTCTGTTC 1986

653 SerMetLeuAsnTyrGluAsgThrLysHisProHisLeuMetGlySerSerValLeuGly 672

1987 TCCGTTCTGAACTACGAAACGTCTGTCTGTCCGGGTCTGTGGGTCTCCGTTCTGGT 2046

673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692

2047 CTGACGACATCCACCGTCTGGCGTACTCTCTGCGTTCCTGCGTTCGTCAGACCCG 2106

693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712

2107 CCGCCGGAACTGACTTCGTTAAAGTTGACATTCACGGTGTCTAGCACCATCCCGCAG 2166

713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 733

2167 GACGCTCTGACCGAAGTTATCGCTTCCATCTCAAA---CCGACGAACACCTACTCGGTT 2223

733 ArgGlnTyrAlaValAlaArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752

2224 CGTCTGTACGCTGTGTTTCAGAAAGCTGTCTACGGTCACTGCTGTGTAAGACTTTCAATCC 2283

753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772

2284 CACGTTTCCACCTGACCGACCTGACGCGTACATGCTCAGTCTGCTGCTCACCTGCAG 2343

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792

2344 GAAACC-----TCCCGCTGCGTGACGCTGTGTTATCGAACAGTCTCTCTCCTGAAC 2397

793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812

2398 GAAGCTTCTCCGCTGTGTCGACGTTCTCTCGCGTTTCATGTGCCACACGCTGTTCGT 2457

813 IleGlyAspAraCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832

2458 ATCCGTTGTAATCCTACGTTCACTGCCAGGATCCCGACAGGTTCCATCTGTCCACC 2517

833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852

2518 CTGCTGTCTCTCTGTCTGTCACGGTGCATCGAAACAAACTCTCTCGCTGGTATCCGTCGT 2577

853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872

2578 GACGCTCTGCTGCTGGTCTGGTTGACACTTCTCTGCTGTTACCCCGCACCTGACCCAC 2637

873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892

2638 GCTAAACCTCTCTCGCTACCTCGTTTCGTGGTGTTCGGAATAGCGTTGGCTGTAAAC 2697

893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912

2698 CTGCGTAAACCGTGTGTTAACTTCCCGTGTGAAGACGAAGCTCTGGGTGTTACCGCTTC 2757

913 TyrGlnLeuProAlaHisCysLeuPheProTyrPheGlyLeuLeuAspThrGlnThr 932

2758 GTTCAGATGCGCGCTCACGGTCTGTTCCCGTGTGGGTCTGCTGCTGACACCCGTAAC 2817

933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952

2818 CTGGAAGTTCACTCGTCCGACTACTCTCTACGCTGTACTCTCATCCGTCGTTCCCTGACC 2877

953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972

2878 TTCAACCGTGGTTCAAAGCTGGTCGTAACATCGCTGTAACACTGCTCGGTGTTCTCGCT 2937

973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992

2938 CTGAAATGCCACTCCCTGTCTCTGGACCTCGAGTTAACTCTCCTGACACCGCTTGACC 2997



```
|||||
361 TCCTATTGCCAAATCTGTTACTGAGCTTTGAGAGGTTCTGGGCTTGGGTTGTTG 420
|||||
141 LeuSerArgValGlyAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
|||||
421 TTGAGAAGAGTTGGTGATGATGTTTGGTTCTATTGTTGGCTAGATGCTGTTGTTG 480
|||||
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnHisCysAla 180
|||||
481 TTGGTTGCTCCATCTTGCTTATCAAGTTTGTGCTCCACCATTTGATCAATGGGTGCT 540
|||||
181 ThrThrAspLeuProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
|||||
541 GCTACTCAAGCTAGACACACCATGCTTCT--GGTCCAGAAGAAGATTGGGT-- 594
|||||
201 AspPheThrAsnLeuArgPheLeuGlnGlnHisLeuSerSerArgGlnGluAlaPro 220
595 -----TGTGAAGAGCTTGGAAATCATCTCTGTGAGAGAGCTGGT 633
|||||
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240
634 GTTCCATTGGGTTGCGAGCTCCAGGTGCTAGAGAGAGGTGGTCTCTAGATCT 693
|||||
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
694 TTGCCATTGCCAAAAGACCAAGAGAGGTGCTGCCAAGACCAAGAAAGAACTCCAGT 753
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp-- 272
754 GGTCAAGGTTCTGGGCTCATCCAGTAGAAGTACAGAGTCCATCTCATAGAGTTTGT 813
|||||
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
814 GTTGTCTTCCAGCTAGA-----CCAGCTGAAGAAGCTACTCTCTTGGAA 858
|||||
293 GlyLysValSerAspLeuSerLeuSer--GlySerValCysCysLeuLysHisLysProSer 311
859 GGTGCTTTGTCTGGTACTAGACATCTCATCTCATCTGTTGGTAGACAACATCATGCTGGT 918
|||||
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro--Phe 330
919 CCACCATCTACTCTAGACACCAAGACCATCGGATACCTCCATGCTCCACCATTTATGCT 978
|||||
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
979 GAACATAAACATTTTGTATTCTCTGCTGAT--AAAGAACAAATTGAGACCATCTTTT 1035
|||||
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyValAlaArgArgLeuValGluLeuPhe 370
1036 TTGTTGCTCTTTGAGACCATCTTTGACTGCTGCTAGAGATTGTTGAACTATTTT 1095
|||||
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
1096 TTGGGTTCTAGACCATGGATGCCAGGTACTCCAGAAGATTGCCAAGATTGCCAAGA 1155
|||||
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
1156 TATTGCCAATGAGACCATTTTGGGAATTTGTTGGATATCTGCTCAATGCTCATAT 1215
|||||
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
1216 GGTGTTTTTGTGAATACTCATTTGCTCCATTGAGAGCTGCT-----GTTACTCCAGCT 1266
|||||
431 Leu-----TCTCCATTGAGAGAGCTGTTGTTATTTGACAACTTTCTTCTTGAAT 1297
1267 GCTGGTGTGTTGCTAGAGAAACCAAGGTTCTGTTGCTGCTCCAGAAGAAGAGAT 1326
|||||
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
1327 ACTGATCCAGAAGATTGGTCAATTTGTTGAGCAACATTTCTTCCATGGCAAGTTAT 1386
|||||
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
|||||
```



RESIT.T 4

```

recent Similarity: 72.94% Conservative: 122
st Local Similarity: 62.36% Mismatches: 260
ery Match: 59.40% Indels: 52
: 3 Gaps: 13

-09-042-460-2 (1-1122) x US-08-974-549A-641 (1-3396)

1 MetThrArgAlaProArgCysProAlaValA:gsrLeuLeuArgSerArgTyrArgGlu 20
1 ATGCCAAGAGCTCCAAAGATGAGAGCTGTTAGATCTTTGTTGAGATCTCACTACAGAA 60
21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
61 GTTGTGCCATTTGCTCTTTCGTAGAGATGGGTCCACAGAGTTGGAGATGGTTCAA 120
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
121 AGAGTGTACCCAGCTCTTCAGAGCTTTGGTGTCTCAATGTTGGTTGTTCCATGG 180
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
181 GAGCTAGACACACACAGCTGCTCCATCTTTCCAGACAGATTTCTTTGTTGAAGAAATTG 240
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
241 GTTGTAGAGTTTGTCAAGATTTGTGAAGAGGTGCTTAGAAGCTTTTGGCTTCGGT 300
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
301 TTGCTTTGTTGACGGTGTAGAGGTGGTCCACAGAGCTTTCACTACTTCTGTAGA 360
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
361 TCTTACTTGTCCAAACACTGTACTGACGCTTTGAGAGGTTCGTGGGTGTTGGT 420
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
421 TTGAGAAGAGTTGGTGACGAGCTTTTGGTTCACTTGTTGGCTAGATGTCTTGTGCT 480
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
481 TTGGTGTCTCATCTTGTGCTTACCAAGTTTGGTGTCCACCAATTTGACCAATTTGGTGT 540
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
541 GCTACTCAAGCTAGACACACACACAGCTTCT---GGTCCAAAGAGAAGATGGGT--- 594
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
595 -----TGTAAGAGAGCTTGGAAACCACTCTGTGTAGAGAAGCTGGT 633
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
634 GTTCCATTTGGTGGTCCAGCTCCAGGTGCTAGAGAAGAGGTGGTGTCTTCTAGATCT 693
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
694 TTGCCATTTGCCAAAGACACCAAGAGAGGTGCTGTCCAGAACCAAGAAAGAACTCCAGTT 753
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
754 GGTCAAAGTTTCTGGGCTCCAGGCTAGAACCTAGAGGTCCATCTGTACAGAGTTTCTGT 813
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
814 GTTGTCTTCCAGCTAGA-----CCAGCTGAGAAGACTACTTCTTTGGAA 858
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
859 GGTGCTTTGTCTGTACTAGACACTCTCTACCCACTCTGTGGTAGACCAACACACCGCTGT 918
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330

```



FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 642:

SEQUENCE CHARACTERISTICS:

LENGTH: 3396 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY:

LOCATION: 1..3396

OTHER INFORMATION: /note="generic" hTERT protein encoding

-08-974-549A-642

Ignment Scores:

ed. No.: 0 Length: 3396

ore: 3505.00 Matches: 719

rcent Similarity: 72.94% Conservative: 122

st Local Similarity: 62.36% Mismatches: 260

ery Match: 59.40% Indels: 52

: 3 Gaps: 13

-09-042-460-2 (1-1122) x US-08-974-549A-642 (1-3396)

```
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
1 ATGCCACGTCGCCACCGTGTGCGTTCGTTCTTTGTTGCGTCTCTCACTACCGTCAA 60
21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgValGln 40
61 GTTTTCGCAATGGCCACCTTCGTTCTGTTGGGTCCACAGGTGGCGTTGGTTCAA 120
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTTP 60
121 CGTGTGTATCCAGCGCCCTTCGTCGCTTGGTTGCCAATGTTTGTGTTGCCATGG 180
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
181 GATGCCCGTCCACCCACCGCCGCCCATCTTTCCGTCAGTTCTTGTGAAAGAAATTG 240
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
241 GTTGGCCGCTTTTGGCAACGTTTGTGTGAACGTGTGTGCCAATAACGTTTGGCTTCGT 300
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
301 TTCCTCTTGTGGATGTCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTTPMetLeuLeu 140
361 TCTTACTTGCACAAACCGTTACCGAGTCCTTGGTGGTTCGTTGGTTCGTTGGTTCGTT 420
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
421 TTGGTTCGTTGGTTCGTTGGTTCGTTGGTTCGTTGGTTCGTTGGTTCGTTGGTTCGTT 480
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
481 TTGTTGGCCCATCTTGTGCTACCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
```

```
QY 181 ThrThrAspIleTTPProSerValSerAlaSerTyrArgProThrArgProValArg 200
Db 541 GCCACCCAAAGCCCGTCCACCCACCGCTCT---GGTCCAGTGTGTTGGGT--- 594
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
Db 595 -----TGTGAACGTGCGTGGAAACCACTCTGTTCTGTAAGCGGT 633
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240
Db 634 GTTCATTGGGTTTGGCAGCCCGGCGCGTCTGCTGCTGCTGCTGCTGCTCTCTCTCT 693
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
Db 694 TTGCCATTGCCAAACAGTCCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 753
QY 260 -----HisArgGlnValLeuProThrProSerArgLysSerTTP--- 272
Db 754 GGTCAAGGTTCTTGGGCCCCACCCAGGTCGTACCGTGTGTCATCTGATCGTGTTCGT 813
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
Db 814 GTTGTTCCTCCAGCCCGT-----CCAGCCGAGAGGCCACCTCTTTGGAA 858
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 859 GGTGCTTGTCTGTGTACCGGTCACTCTCACTCTCTGTCGTCGTCACACACCGCGGT 918
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 919 CCACCATCTACTCTGTCACACCGTCATGGATACCCATGTCACACCGTTTACGCC 978
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 979 GAAACCAACACTTCTTGTACTCTTCTGTGTAT---AAGAACAATGTCGTCATCTTTC 1035
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgGluValGluIlePhe 370
Db 1036 TTGTGTCTTCTTGTGTCATCTTGTACCGGTGCGGTGTTGTTGTTGAAACCAATTTTC 1095
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1096 TTGGGTTCTCGTCCATGATGCCAGGTACCCACGTCGTTGCCACGTTTGCACACGT 1155
QY 391 TyrTTPGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
Db 1156 TACTGGCAATGCGTCCATCTTCTTGGAAATGTTGGGTAAACACGCCCAATGTCCATAC 1215
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
Db 1216 GGTGTTTGTGAAACCCACTGTCCATGTCGTCGTCGCC---GTTACCCAGCC 1266
QY 431 Leu-----Asn 432
Db 1267 GCGCGTGTGTGCGCGGTAACCAAGGTTCTGTTGCCGCCCCAGAGAAGAAGAT 1326
QY 433 ThrSerProProHisLeuMetAspLeuArgLeuHisSerSerProTTPGlnValTyr 452
Db 1327 ACCGATCCACGTCGTTGGTTCAATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1386
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTTPGlyThrArgHis 472
Db 1387 GGTTCGTTGCTGCTGTTTGTGTCGTCGTTTGTGTCACACGAGTTTGGGTTTCTCGTCAC 1446
QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
Db 1447 AACGAACGTCGTTTCTTCGTCGTAACCAACCAAAATTCATTTCTTTGGGTAAACACGCCAA 1506
QY 493 LeuSerLeuGlnGluLeuMetTTPLysMetLysValGluAspCysHisTTPLeuArgSer 512
Db 1507 TTGTTTGTGCAAGAAATGACCTGGAATAATGTCGTTGTAATGTGTCGTTGCGTCGT 1566
```

513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 TCTCCAGGTGTTGGTGTTCAGCGCGGACACCGGTTGCGTGAAGAAATTTGGCC 1626  
533 ThrPheLeuPheThrLeuMetAspThrTyValValGlnLeuArgSerPhePheTy 552  
1627 AATCTTGCACCTGGTGTGATCTGTACGTTGTGAATGTTGGTCTTCTTCTAC 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyArgLysSerValTrpSer 572  
1687 GTTACCGAAACCACTTCCAAAAAACCGGTTGTCTTCTACCGTAAATCTGTTGGTCT 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AATTCGAACTATGTGATTCGTCACACTTGAACGTTCAATTCGGTGAATGTCT 1806  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAGCCGAAGTGTGTCAACACCGTGAAGCCGTCACGCTTGTGACCTCTCGTTGGGT 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTySerMetGlyThr 632  
1867 TCAATTCAAAACAGATGTTGCGTCCAAATGTAAATGATGATGATGATGATGATG 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 CGTACCTTCGTGTGAAAAACGTCGCGACGTTTGCACCTCTCGTGTAAAGCCTGTTC 1986  
653 SerMetLeuAsnTyGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 TCTGTTTGAACACTACGACGTCGCGTCCAGGTTGTGGGTGCGCTCTGTTTGGT 2046  
673 MetAsnAspIleTyArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 TTGGATGATATTACCGTGCCTGCGGACCTTGTGTTGGTGTTCGTCGCCAAGATCCA 2106  
693 ThrProArgMetTyPheValLysAlaAspValThrGlyAlaTyAspAlaIleProGln 712  
2107 CCACAGAAATGTACTTGTGTTAAAGTTGATGTACCGGTGCTTACGATACCATTCACAAA 2166  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyCysIle 732  
2167 GATCGTTTGCAGGATATTGCTCTATTATTAAA---CCACAAACACCTACTGTGTT 2223  
733 ArgGlnTyAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 CGTCGTTACGCGGTTGTTCAAAAAACGCGCCACCGTCACTGCTGTAAGCCTTCAATCT 2283  
753 GlnValThrLeuSerAspLeuGlnProTyMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CACGTTTCTACCTTGACCGATTTGCACCATACATGCGTCAATTCGTTGCCACTTGCAA 2343  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2344 GAAACC-----TCTCCATTGCGTGATGCGGTTGTTATTGAACAACTCTTCTTTGAAC 2397  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAAGCCTCTCTGTTGTTGATGTTTCTTGGCTTCATGTTCCACGCGGTTGCT 2457  
813 IleGlyAspArgCysTyThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATTCTGGTGAATCTTACGTTCAATGTCAAGGTATCCACAAAGGTTCTATTGTTGTCTACC 2517  
833 LeuLeuCysSerLeuCysPheCysPheCysMetLysLeuPheAlaGluValGlnArg 852  
2518 TTGTGTGTCTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2577  
853 AspGlyLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872  
2578 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 926  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyArgCysMetIleAsn 892

2638 GCCAAAACTTCTTGGCGTACCTTGTTCGTTGGTGTTCAGAAATACGTTGTGTGTTAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGTAAACCGTGTGTTAACTTCCAGTTGAAGATGAAGCCCTTGGGTGTACGCGCTTC 2757  
913 TyGlnLeuProAlaHisCysLeuPheProTyCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAATGCGCAGCCACCGGTTGTTCCCATGGTGTGTTGTTGTTGTTGTTGTTGTTGTT 2817  
933 LeuGluValPheCysAspTySerGlyTyAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 TTGGAAGTTCAATCTGATTAATCTTCTTACGCGCTACCTCTATTGCGCTCTCTTACCC 2877  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGTGGTTTCAAGCCGCTGCTAACATGCTGCTGAATGTTGTCGGTGTGTTGCGT 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 TTGAATGTCACCTCTTGTCTTGGATTGCAAGTTAACTCTTTCGAAACCGTTGTGACC 2997  
993 AsnIleTyLysIlePheLeuLeuGlnAlaTyArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATTTACAAATTTGTTGTGCAAGCCTACCGTTTCCACGCTGTGTTTGCATTTG 3057  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTCACCAACAAAGTTTGGAAAAACCAACCTTCTTTCGCGTATTATTCTGTATACC 3117  
1033 AlaSerCysCysTyAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCTTGTGTACTTATTTTGAAGCAAAACGCGGTATGCTTTCGGTGCCAAA 3177  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyArgGlnAlaPheLeu 1069  
3178 GGTGCGCGCGTCCATTGCCATCTGAAGCGGTCAATGTTGTGTCACCAAGCCTTCTTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 TTGAATTCACCGCTCACCGTGTATCTACCTACGTTCCATGTTGGTGTCTTGGTACCGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3298 CAACCAATGTCTGTAATTCAGTACCATCCTTGACCGCTTGGAGCGCGCCGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCCAGCCTTGCCATCTGATTTTCAAAACCAATTTGGAT 3396

RESULT 6  
US-08-974-549A-721  
; Sequence 721, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lindner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20

28 ATGCGGGTGTCTCCGGTTCCTGCTGTTTCCTGCTGCCTATCCCATATCGCGAA 87

21 ValTrpProLeuAlaThrPheValAArgLeuGlyProGluGlyArgArgLeuValGln 40  
:::  
:::

88 GTTCTGCCGTGGCTACTTCGTTCTGCTCTGGCCCGCAGGGTGGCGTCTGGTCTCAG 147

41 ProGlyAspProLysIleTyArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
:::  
:::

148 CGTGTAACCGGCTGCTTCCTGCTCTGTTGCTCAGTGCCTGGTTTGGTTCCGTGG 207

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
::::  
::::

208 GACGCTGCTCCGCCCGGCTGCTCCGTCCTTCCTCCGTCAGGTTTCCTGCTCAAAAGAATG 267

81 VallalaArgValValGlnAArgLeuCysGluAArgAsnGluAArgAsnValLeuAlaPheGly 100  
:::  
:::

268 GTTGCTGTGTTCTGACGCTCTGTCGAACGTGGTCTAAAACAGTTCTGGCTTCGGT 327

101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
:::  
:::

328 TTCGCTCTGCTGACGGTGTCTGTTGGTTCGCGGAAGCATTCACCACCTCCGTCGT 387

121 SerTyrLeuProAsnThrValIleGluThrLeuAArgValSerGlyValAArgTrpMetLeuLeu 140  
:::  
:::

388 TCCCTACCTGCCGAACACCGTTACCGACGCTCTCGGTGGTTCGGTGCCTGGGGTCTGCTG 447

141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
:::  
:::

448 CTGGCTGTGTTGGTAGCAGCTTGTGTTACCTGCTGGCTCGTGGCTCTGTTGCTGT 507

161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
:::  
:::

508 CTGGTTGCTCGCTCTCGCTTACCAAGTTTGTGGTTCGCGCTGTACAGCTGGGTGCT 567

181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
:::  
:::

568 GCTACCCAGGCTGCTCGCCGCCGACGCTTCC---GGTCCGCTGCTGCTCGGT--- 621

201 AsnPheThrAsnLeuAArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
:::  
:::

622 -----TGGAAACGTGCTTGGAACCACTCCGTCGCTGAAGCTCGT 660

221 LysProLeuAlaLeuProSerArgGlyThrLysAArgHisLeuSerLeuThrSerThrSer 240  
:::  
:::

661 GTTCGCTGGGTCTCGCGGTCCGCGTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720

241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
:::  
:::

721 CTGCGCTGCCGAACGTCGCGTCTGGGTGCTGCTCCGMAACCGGAACGTAACCCCGGT 780

260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
|||  
|||

781 GGTACAGGTCTCTGGGTCTACCCGGTCTGTACCGGTGGTCCGTCGCAACCGTGGTTCTGC 840

273 ValProSerProAlaAArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
|||  
|||

841 GTTGTCTCCCGGCTCT-----CCGCTGAGAAGACTACCTCCCTCGGNA 885

293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
|||  
|||

886 GGTGCTCTTCGTCGCGCACCGCTCACTCCACCGCTCGTTGGTTCGTGCAACCACTCGT 945

312 SerThrSerLeuLeuSerProProAArgGlnAsnAlaPheGlnLeuAArgPro---PheIle 330  
|||  
|||

946 CCGCGCTCCACTCCGTCGCGCGCTCGGTGGACAACCCCGTGGCCCGCGGTATTACGCT 1008

331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
:::  
:::

1006 GAACAACAACACTTCTCTGTAATCTCTCCCGGTGAC---AAGAAGACGTGGTCCGCTTC 1062

351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyValAArgArgLeuValGluIleIlePhe 370  
:::  
:::

1063 CTGCTGTCTCTCCGCTCGCTCGCTGACCGGTGCTCGTCTGGTGGACTATCTTC 1122



371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1123 CTGGGTTCCCGTCGTCGATCGCGGACCCCGCGTCGTCGCGCGTCGTCGCGAGCGT 1182  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGlnCysGlnTyr 410  
1183 TACTGGCAGATGCGTCGCTGTTCTCTGGAATGCTGGCAACACCGCTCAGTGGCCGAT 1242  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1243 GGTGTCTGCTGAAACCCACCTGCGCGCTCGCTGCT-----GTACCCTCGCT 1293  
431 Leu-----Asn 432  
1294 GCTGGTGTTCGCTGCGTGAACCCGACGGTTCGCTGCTCCGGAAGAGAT 1353  
433 ThrSerProProHisLeuMetAspLeuLeuArgHisSerSerProTrpGlnValTyr 452  
1354 ACCGACCGCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1414 GGTTCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheLeuLeuLysLysLysLys 492  
1474 AACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1534 CTGTCCTGCGAGAACTGACCTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593  
513 SerProGlyLysAspArgValProAlaAlaGlnHisArgLeuArgGluArgIleLeuAla 532  
1594 TCTCCGGGTGTTGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1654 AAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1713  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrpSer 572  
1714 GTTACCGAAACCCCTCCGAAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1773  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1774 AAATGTCAGTCCATCGGTATCCGTCAGCACCTGAAACGCTGTTTCAGCTGCTGAACTGCC 1833  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
1834 GAAGCTGAAGTTCGTCAGCACCGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1893  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1894 TTTCATCCGAAACCGACGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1953  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1954 CGTACCTTCCGTCGTGAACCGTCTGAGCGCTGACCTCCCGCTGTTAAAGCTCTGTT 2013  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2014 TCGTTCGTAATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2074 CTGACGACATCCACCGTCTGGGCTACCTTCGTTCTGCTGCTGCTGCTGCTGCTGCT 2133  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2134 CCGCCGGAACCTACTCTGTTAAAGTTACCGTTACCGCGCGGTACGACACCATCCCGCAG 2193  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2194 GACCGTCTGACGAGGTATCGCTTCCATCATCAAA-----CCGCAAGACCATCTATGCTT 2250  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2251 CGTCGTTACGCTGTTGTTTCAGAAAGCTGCTCAGCGTCACTGCTAAAGCATTCAAATCC 2310  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisIleGln 772  
2311 CAGGTTTCCACCTGACCGCTGACCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 2370  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2371 GAAACC-----TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2424  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2425 GAAGCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2484  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2485 ATCCGTTGTAATCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2544  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2545 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2604  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2605 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2664  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2665 GCTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2724  
893 LeuGlnLysThrValValAsnPheProGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2725 CTGCGTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2784  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2785 GTTCAGATGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2844  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2845 CTGGAAGTTCAGTCCGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2904  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2905 TTCAACCGTGTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2964  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2965 CTGAAATGCGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3024  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3025 AACATCTACAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3084  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3085 CCGTTCACCGACGAGTTTGGAAAAACCGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 3144  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3145 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3204  
1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3205 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3264  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089

3265 CTGAACTGACCGTCCCGTGTACCTACGTTCCGTCGCTGGGTTCCCTCGGTACCGCT 3324  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3325 CAGACCCAGCTGTCCGCTAACTGCGGGTACACCCCTGACCGCTCTGGAAGCTGCTGCT 3384  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3385 AACCCGCTCTGCCGTCGCACCTTCAAACCATCTCGGAC 3423

SULT 7  
-08-851-843A-224  
Sequence 224, Application US/08851843A  
Patent No. 6093809  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6093809el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESS/SEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
OTHER INFORMATION: /note= "human telomerase reverse  
transcriptase (hTERT) catalytic protein

OTHER INFORMATION: component "  
US-08-851-843A-224  
Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 13 Gaps: 13  
US-09-042-460-2 (1-1122) x US-08-851-843A-224 (1-4015)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
DB 56 ATGCCGGCGCTCCCGCTGCCGAGCGCTGCCCTCTGCTGCCGACCACTACCGCGAG 115  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
DB 116 GTGCTGCCGCTGCCACAGTTGTCGCGCGCTGCCGCGCCCGCGGCTGCCGCTGTCAG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
DB 176 CCGGGGACCCCGCGCTTTCGCGCGCTGTCGCCAGTGCCTGCTGTCGCTGCCCTGG 235  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
DB 236 GACGACCGCGCGCCCGCGCTTCTTCCGCGAGGTGCTCCTGCTGAAGAGAGCTG 295  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
DB 296 GTGGCCCGAGTGTCTCAGAGGCTGTCGAGCGCGCGCGCGAGAACGTCGCTTGGC 355  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
DB 356 TTCGCGCTGCTGACCG 415  
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
DB 416 AGCTACCTGCCCAACACCGGTGACCGACCTGCGGGGAGCGGGGCGTGGGGCTCTG 475  
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
DB 476 CTGGCGCGGTGGGCGAGCGAGCTGCTGCTTCCCTGCTGGCACCTGCGCGCTCTTTGTG 535  
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
DB 536 CTGGTGGCTCCACAGCTGCCCTTACCAGGTGTGCGCGCGCGCGCTGTACCACTCGCGCT 595  
QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
DB 596 GCCACTCAGCG 649  
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
DB 650 -----TGCAGACCGCGCTGGAACCACTAGCTCAGCGGAGCGCGCGG 688  
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
DB 689 GTCCCCCTGGGCTGCCAGCCCGCGGTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGAGT 748  
QY 241 ValProSerAlaLysLeuAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
DB 749 CTGCGCTTGCACAGAGCTT 808  
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
DB 809 GGCAGGGGTCTCTGGGCCCG 869  
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
DB 869 GTGGTCTCAGCTGCCAG-----CCCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCG 913

293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GGTGGCTCTCTGGCAGCGCGCACTCCCAACCCATCCGTGGCGCGCCAGCACACGCGGCG 973  
312 SerThrSerLeuLeuSerProProAsnGlnAsnAlaPheGlnLeuArgPro---Phe11e 330  
974 CCCCACATCCAGCGCGCACACAGTCCCTGGACAGCGCTTGTCCCGGTGTACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTCTCTACTCTCTCAGGGGAC---AAGGAGCAGCTGGCGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeuPhe 370  
1091 CTACTCAGCTCTCTGAGCGCCAGCTGACTGCGCGCTCGAGGCTCTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGGTTCCAGGCGCTGGATCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCGCTCTCTGAGGTGCTTGGAGTCTCTGGAAACACGCGAGTCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTCG---GTACCCCGCAGCA 1321  
431 Leu-----Asn 432  
1322 CGCGGTGTCTGTCGCGGAGAACCCCGAGGCTCTGTGGCGGCCCCGAGGAGGAGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGGCTGGTGCAGCTGCTCGCCAGCACAGCAGCCCTCGCAGGTGTAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGTCGCGCGCTGCTGCGCGGTGTGTGCCCGCCAGGCTCTGGGCTCCAGCGAC 1501  
473 AsnGluArgArgPhePheAsnLeuLysLysPheLeuSerLeuGlyLysTyrGlyLys 492  
1502 AACGACCGCGCTTCTCAGAACACCAAGAGTTTCATCTCCCTGGGAGAGATGCCAAG 1561  
493 LeuSerLeuGlnGluMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTGCTCAGGAGCTGACGTGGAAGATGAGCGTGGGACTCGCTTGGCTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlyLeuAla 532  
1622 AGCCAGGGGTGGCTGTCTCCGCGCCGAGACACCGCTCTGCGGTGAGGAGATCTTGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTCTCTGCTGATGATGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACAGGAGACCAAGCTTTCAGAAAGACAGGCTCTTTTCTACCGAAGAGGTCTGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTCGAAGCATTTGATCAGACAGCATTTGAAGAGGTGCAGCTCGGAGCTGTGCG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTGAGGATCGGAAAGCAGGCGCCCTGCTGACGTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCGAGCTGACGGGTGGCGGATTTGACATGGACTGACTGCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652

1982 AGAACCTTCGCGAGAAAGAGGCGCGAGCTCTCACCTCAGGGTGAAGGCACTGTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTCTCAACTACGAGCGCGCGCGCGCGCTCTCTGGGCGCTCTGTGTGTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGAGCATATCCAGGCGCTGGCGCACCTTCGTGCTGCGGTGCGGGCCAGGACCGC 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGACTTGTCAAGGTGATGTGACGGCGGTACGACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGTCAAGGAGTATCCAGCATCAAA---CCCCAGAACACGACTGCGGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCCAGAGGCGCGCCATGGCGACGTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTCTACCTTGACAGCTCCAGCGGTATCGGACAGTTCGTGGCTCACCCTGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGTCTCATCGACGAGCTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCCAGCAGTGGCTCTTCGACGTCTTCACGCTTCATGTGCCACACCGCGCTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGCGAAGTCTTACGTCCAGTCCAGGAGATCCCGCAGGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACGCTGTCTACGCGCAGCATGGAGAACAGCTGTGTGCGGGGATTCGCGCG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 CGCAAAACCTCTCTCAGGACCTTGGTCCGAGGTGTCTCTGAGTATGGCTGCTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaPro 912  
2753 TTGCGAAGACAGTGGTGAACCTTCCTCTGATAGACAGCGGCCCTGGGTGGCGACTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGCCACCGCTATTCCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGACAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGACGCGACTCTCAC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGCTTCAAGGTGGAGAACATGCTGCTGCAAACTCTTTGGGTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCGCTTCTGAGTTTTCAGGTGAACAGCGCTCCAGACGCGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012



141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 CTGCCCGCGTGGCGAGCGAGCTGTCTTACCTGTGGCGACCTGGCGCTTTTGTG 535  
y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla 180  
b 536 CTGGTGGCTCCAGCTGGCTTACAGGTGTGGCGCGCTGTACCTGACCTGGCGCT 595  
y 181 ThrThrAspIleTyrProSerValSerValSerTyrArgProThrArgProValGlyArg 200  
b 596 GGCACCTCAGCGCGCGCGCGCACACACTAGT---GGACCCCAAGCGCTCTGGGA--- 649  
y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
b 650 -----TGCAACGGCGCTGGAACCATACGTCAGGGAGCGCGG 688  
y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
b 689 GTCCCCCTGGGCTGCGAGCGCGCGGTGCGAGGAGCGCGCGGCGAGTGCACCGAAGT 748  
y 241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTGCCCTTCCCAAGAGCCAGCGGTGGCGCTGCCCTGAGCCGGAGCGGACGCCGTT 808  
y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
b 809 GGGCAGGGCTCTGGGCGCCACCCGGCGAGGACGCGTGGACCGGTGCTGTTCTGT 868  
y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTACCTGCCAGA-----CCCGCCGAAGAGCGACACTCTTTGGAG 913  
y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GTGGCGCTCTGGCAGCGCGCACTCCACCCATCCGTGGCGCCGACCCAGCGCGGC 973  
y 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCCCACATCCAGCGCGCACAGCTCCCTGGGACAGCGCTTGTCCCGCGGTGTACGCC 1033  
y 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGAC---AAGGAGAGCTGCGGCCCTCTTC 1090  
y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTGAGGCCAGCGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTT 1150  
y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
b 1151 CTGGGTTCAGGCGCTGGATGCCAGGACTCCCGCAGGTTGCCCGCGCTGCCCGACGC 1210  
y 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
b 1211 TACTGGCAATCGGCGCTGTTCTTGAGCTGCTTGGGAACCCAGCGAGTGCCTTAC 1270  
y 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
b 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGGCGAGTGGC-----GTCACCCACGCA 1321  
y 431 Leu-----Asn 432  
b 1322 GCGGGTGTGTGCCCGGAGAACCCAGGCGCTCTGTGGCGCGCCCGAGGAGGAGGAC 1381  
y 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
b 1382 ACAGACCCCGCTGCGCTGGTGTGCTGCTCCCGCAGCAGCAGCGCCCTGGCAGGTATC 1441  
y 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472  
b 1442 GGCTTGTGGCGGCTGTCTGCGCGGTGTGGTCCCGCAGGCGCTCTGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492  
1502 AACCAACCGCGCTTCTCAGAAACACCAAGAGATTCATCTCTCTGGGAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGCTCAGGAGCTGAGTGGAAAGATGAGCGTGGGAGCTGGCTGGTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCACGGGTGGCTGTCTCCGGCGCAGACACCTCTTCTTACCGAAGAGTGTCTGGAGC 1681  
533 ThrPheLeuPheTyrPheLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCTGCTGCTGCTGATGAGTGTGTACGTCGTGAGCTGTCTCAGGTCTTCTTTAT 1741  
553 IleThrGlnSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrPhe 572  
1742 GTCCGCGAGACCGTTCAAAGAAACAGAGCTCTTTTCTTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTCAAGAGGCTGAGCTGGCGAGCTGTGC 1861  
593 GlnGluGlnValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGCTCAGGAGCATCGGAGAGCAGCGCCCTGTGTGACGTCAGACTCCAGCTCC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTATCCCAAGCTGTACGGCTGTGGCGCATGTGAAATGAGTACCTGCTGGGAGGC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGGCGCGAGCGCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuLeuTyrGlnArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGCGCGCGCGCGCTCTGGCGCCCTCTGTGTGGGCG 2101  
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGAGCATATCCACAGGCTGGCGCACCTCTGTGTGTGGCGGCCAGGAGGCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGGTGATGTACCGGCGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCAGCATCATCAA---CCCGAAGACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGATGCGGTGTCCAGAGCGCGCCATCGGCGCGCTCCGCAAGCGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTCTACTTTCAGACAGCTCCAGCGCTACATGCGACAGTTCGTGCTCACTGCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACCC---AGCCGCTGAGGATGCGCTGTCTATCGACGAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheHisPheLeuArgHisSerValValLys 812  
2453 GAGCGACAGTGGCGCTCTCGAGCTTCTTACGCTTCTGAGTTCAGCCACCGCGCTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTCAGTCCAGTCCAGGGATCCCGAGGGCTCCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuLeuPheGlyAspMetGluLysLeuPheAlaGluValGlnArg 852

```
2573 CTGCTCTGAGCTGTGCTACGGCGACATGGAGAAAGCTGTTCGCGGATTCGCGG 2632
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872
2633 GACGGCTCTCTCTGCTGTGTGTGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2692
873 AlaLeuThrPheLeuSerThrLeuValHisGlyValProGluThrGlyCysMetIleasn 892
2693 GCGAAACCTTCTCTAGGACCTGTGTCAGGTGTCCTGAGTATGGCTGGTGTGAAC 2752
893 LeuGlnLeuThrValValAsnPheProValGluProGlyThrLeuGlyGlyValAlaAlaPro 912
2753 TTGCGGAGACAGTGTGTGACTTCTCTGTAGAGAGAGGCGCTGCTGGTGGCAGCGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProThrCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATCGCGGCCACCGCTATTCTCCCTGTGTGCGGCTGTCTGTGATACCGGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLeuThrSerLeuThr 952
2873 CTGAGGTGACAGCGACATCTCCAGCTATGCGGACCTTCATCAGAGCCAGTCTCACC 2932
953 PheGlnSerValPheLeuValGlySerGlyThrMetArgAsnLysLeuLeuSerValLeuArg 972
2933 TTCACCGCGGCTTCAAGCTGGAGGACATGCGTGGCAAACTCTTTGGGTCTTGGCG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGCTGTCTTCTGATTTGCGAGTGAACAGCTCCACAGCGTGTGACC 3052
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAAGATCTCTCTGCTGCGAGGCGTACAGTTTCAGCATGTGTCTGCGCTC 3112
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTCGCGTCTCATCTCTGACACG 3172
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCCTCCCTCTGCTACTCCATCTCTGAGGACCAAGACGAGGATGCTCGTGGGGCCAG 3232
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGGCGCGCGCGCTCTGCGCTCCGAGGCGTGCAGTGTGCTGTGCCACCAAGCATTCCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAAGTGTACTGACACCGTGTCACTACGTGCGCACTCTGGGTCTACTCAGGACAGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
3353 CAGACGCGAGTGTAGTCCGAAGCTCCCGGAGCGACGCTGACTGCTCGCGCGCGAGCC 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCGGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3451
```

SULT 9

-08-854-050-224

Sequence 224, Application US/08854050

Patent No. 6261836

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Haxley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6261836el Telomerase

NUMBER OF SEQUENCES: 225

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 224:
SEQUENCE CHARACTERISTICS:
LENGTH: 4015 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 56..3454
OTHER INFORMATION: /product= "hTERT"
OTHER INFORMATION: /note= "human telomerase reverse
transcriptase (hTERT) catalytic protein
component"
US-08-854-050-224

Alignment Scores:
Pred. No.: 0 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
Query Match: 59.40% Indels: 52
DB: 3 Gaps: 13

US-09-042-460-2 (1-1122) x US-08-854-050-224 (1-4015)

Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGCGCTCCCTGCTGCGCAGCCACTACCGGAG 115
Qy 21 ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGCTGGCCACGTTCTGTGCGGCGCTTGTGGGCGCCCGAGGCTGGCGCTGGTGCAG 175
```



41 ProGlyAspProIysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
176 CGCGGGACCCGGCGGCTTTCCGGCGGCTGGTGGCCAGTGGCTGGTGGCTGGCTGG 235  
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
236 GACGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 295  
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
296 GTGGCCCGAGTGTGCAGAGGCTGTGCAGCGCGGCGGCGGCGGCGGCGGCGGCGG 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerValArg 120  
356 TTCGCGTGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 415  
121 SerTyrLeuProAnthrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
416 AGCTACCTGCCACACAGTGCACCGACCTGCGGGGAGCGGCGGCGGCGGCGGCGG 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGCTGGCTACACAGTGTGGGCGCGCGCTGTACAGCTGGCGCT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCG 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGCACGGCGCTGGACCATAGCTCAGGAGCGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCGCTGGCGCTGCCACCGCGGTCGAGAGGCGCGGCGCGCGCGCGCGCGCG 748  
241 ValProSerAlaLysLeuAlaArgCysTyrProValProArgValGluGluPro 259  
749 CTGGCGTTGCCAGAGGCCAGCGCGTGGCGCTGCCCTCAGCGGAGCGGCGCGCGG 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
809 GGGCAGGGGTCTCGGGCCACCGCGGACGAGCGGTGGACCGAGTACCGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTACCTGCCAGA-----CCCGCGGAGAGGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysLysHisLysProSer 311  
914 GTGGCGCTCTGGCAGCGGCACCTCCACCCATCCGTGGGCGCGCAGCAGCGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCACATCGCGGCGCACCGTCCCTGGGACACGCGCTTCCCGCGGTGTACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTTACTCTCTCAGCGAC---AAGGAGCAGCTCGCGCTCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCCGACGCTGACTGGCGCTCGGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCCAGGCGCTTGGATGCCAGGAGTCCCGCGAGGTGGCGCGCGCTGCCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410

Db 1211 TACTGCAAAATCGGCGCGCTGTTCTGGAGCTGCTGGGAACACGCGCGAGTCCCTAC 1270  
Qy 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
Db 1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGTGGG-----GTCAACCCAGCA 1321  
Qy 431 Leu-----Asn 432  
Db 1322 GCGGTGCTGTGCGCGGAGAGAGCCCGAGGCTCTGTGGCGGCGCGCGAGGAGGAC 1381  
Qy 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
Db 1382 ACAGACCCCGCTGCTGGTGCAGCTGCTCGGCAGCAGCAGCAGCCCTGGCAGGTGTAC 1441  
Qy 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTCTGCGCGGCTGCTGCGCGGCTGGTGGCCCGAGGCTCTGGGGCTCCAGGGAC 1501  
Qy 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
Db 1502 AACGAACGCGCTTCTCAGGAACACCAAGAGTTCACTCCCTGGGGAAGCATGCCAAG 1561  
Qy 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTGCGAGAGTGCAGTGCAGATGAGCGTGCAGGCTGGCTGGCTGGCGCAGG 1621  
Qy 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCAGGGGTGGCTGTGTTCCGCGCGCAGCAGCAGCTCTGCGTGGAGAGATCTGGCC 1681  
Qy 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1682 AAGTTCCTGCACTGGCTGATGATGTGTACGTGTGTGTGTGTGTGTGTGTGTGTGT 1741  
Qy 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTCAAGGAGACCACTGTTCAAAGAACAGGCTCTTTCTACCGGAAGAGTGTCTGGAGC 1801  
Qy 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgLysLeuSer 592  
Db 1802 AAGTTGCARAGCATTTGAATCAGACAGCATTTGAAGAGGTCAGTGCAGGAGTGTCTG 1861  
Qy 593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGTCCAGGAGCATCGGGAAGCAGCGCGCGCTGTGTGTGTGTGTGTGTGT 1921  
Qy 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTTATCCCGAGCTCAGCGGCTGCGCGCGCTGCGCGCTGTGTGTGTGTGTGTGTGT 1981  
Qy 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAACGTTCCGACAGAAAGAGGCGCGCTCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
Qy 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCAACTACAGCGCGCGCGCGCGCGCTCTGTGTGTGTGTGTGTGTGTGT 2101  
Qy 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGACCATATCCAGAGGCGCTGGCGCACCTTGTGTGTGTGTGTGTGTGTGTGTGT 2161  
Qy 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
Db 2162 CCGCTGAGT 2221  
Qy 713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGGCTCAGGAGGTCATCGCAGCATCATCAAA---CCCCAGAACACGTACTCGGTG 2278  
Qy 733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752

Alignment Scores:	
Pred. No.:	0
Score:	3505.00
Length:	4015
Matches:	719

Percent Similarity: 72.94% Conservative: 122  
1st Local Similarity: 62.36% Mismatches: 260  
Very Match: 59.40% Indels: 52  
3: 4 Gaps: 13

3-09-042-460-2 (1-1122) x US-09-430-323-224 (1-4015)

1 MetThrArgAlaProAArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCTCCCTGCTGCGCGACCTACCGCGAG 115  
21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluArgArgValGln 40  
116 GTGTGCGCTGCGACGCTTCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCG 175  
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
176 CGCGGGGACCGCGCGCTTCCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCG 235  
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyLeu 80  
236 GACGACG 295  
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
296 GTGGCGCGCGCTGCGAGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
356 TTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140  
416 AGCTACTGCG 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 GTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGTGTGCTCCCGAGCTGCGCTACAGAGTGTGCGCGCGCGCGCGCGCGCG 595  
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGCG 649  
201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGGGACCGCGCGCTGGACCACTAGCGTCAGGAGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCGCTGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro 259  
749 GTCCCGTTCGCCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr 272  
809 GGGCAGGGGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTCACCTGCCAGA-----CCCGCGGAGGAGGCGCGCGCGCG 913  
293 GlyLysValSerAspLeuSerLeuSer--GlySerValCysCysLysHisLysProSer 311  
914 GTTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro--PheIle 330

974 CCCCCATCCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACCTTCTCTACTCTCTCAGCGGAC---AAGGAGCAGCTCGCGCGCTCTCTTC 1090  
351 LeuLeuSerAsnLeuLysProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGTTCACGCGCGCTGGATGCCAGGACTCCCGCGAGTTGCCCGCGCTGCCCGAGCGC 1210  
391 TyrTyrGlnMetArgProPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATCGCGCGCGCTGTTCTGGAGCTGCTTGGGAACACCGCGAGTCCCGCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGAGCGACTGCCCGCTCGAGCTGCG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCGTGCTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGCGCGCGCGCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452  
1382 ACAGCCCCCGCTGCTGTGTGCGAGCTGCTCCGCGACGACGAGCGCGCTGGCAGGTGTAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472  
1442 GGCTTCGTGCGCGCGCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCG 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492  
1502 AACGAACCGCGCTCTCTCAGGAACACCAAGAAAGTTCTCTCTCGGAGAGATGCCAAG 1561  
493 LeuSerLeuGlnLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGCTGCGAGCGCTGACGTGGAAGATGAGCGTGGGAGCTGGCGCTTGGTGGCGAG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCACGCGGTGCTGTGTCTCGCGCGCGAGAGACCGTCTGCGTGGAGAGATCTCTGGCC 1681  
533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1682 AAGTTCTCTGCGCTGCTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572  
1742 GTACGAGGACCGCTTCAAAAGAACAGCGCTCTTTTCTACCGGAGAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGAATCAGACGACCTTGAAGAGGCTGCGAGCTGCGGAGCTGTG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTyrLeuAlaMetProLysCysArgLeuArg 612  
1862 GAAGCAGAGGTTCAGGAGCATCGGGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1921  
613 PheIleProLysProLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTTCAGCGCGCTGCGCGCGATTTGGAACATGAGTACTGCTGGGAGCG 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGAGAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCTCACTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2101

QY	1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer	1051
DB	3173	GCCTCCCTCTGCTACTCCATCTCTGAAGAACGACGAGGATGTCTGCTGGGGCCCAAG	3232
QY	1053	GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu	1069
DB	3233	GGCGGCGCGCCCTCTGCCTCCGAGGCGGTGCAGTGGCTGTGGCACCAAGCATTCCTG	3292
QY	1070	LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla	1089
DB	3293	CTCAAGCTGACTGCACACCCGTGTCACTACGTGCACACTCTCTGGGGTCACTCAGGACGCC	3352
QY	1090	GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla	1109
DB	3353	CAGACGAGCTGAGTCCGAGAGCTCCCGGGACGACGCTGACTGCCCTGGAGCGCCGACGCC	3412
QY	1110	AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp	1122
DB	3413	AACCGGCACTGCCCTCAGACTTCAAGACCATCTGGAC	3451
RESULT 11			
US-09-572-423B-3			
; Sequence 3, Application US/09572423B			
; Patent No. 6331399			
; GENERAL INFORMATION:			
; APPLICANT: Brett P. Monia			
; APPLICANT: William A. Gaarde			
; APPLICANT: Edward Wanciewicz			
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION			
; FILE REFERENCE: ISPH-0462			
; CURRENT APPLICATION NUMBER: US/09/572,423B			
; CURRENT FILING DATE: 2000-05-16			
; NUMBER OF SEQ ID NOS: 29			
; SEQ ID NO 3			
; LENGTH: 4015			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (56)...(3454)			
US-09-572-423B-3			
Alignment Scores:			
Pred. No.: 0 Length: 4015			
Score: 3505.00 Matches: 719			
Percent Similarity: 72.94% Conservative: 122			
Best Local Similarity: 62.16% Mismatches: 260			
Query Match: 59.40% Indels: 52			
DB: 4 Gaps: 13			
US-09-042-460-2 (1-1122) x US-09-572-423B-3 (1-4015)			
QY	1	MetThrArgAlaProAcGysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu	20
DB	56	ARGCGCGCGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGGCAGCCACTACCGGAG	115
QY	21	ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln	40
DB	116	GTGCTCGCGCTGGCAGCGTTCGTGCGCGCGCTCGGGCCCGCGGCTGGGTGGTGGAG	175
QY	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr	60
DB	176	CGCGGGACCCGGCGGCTTTCCGGCGCCTGGTGGCCAGTGGCTGGTGGTGGCCCTGG	235
QY	61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
DB	236	GACGACGACGCGCGCGCGCGCCCTCCCTCCGCGCAGTGTCTGCTGCTGAAGGAGCTG	295
QY	81	ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
DB	296	GTGGCCCGAGTCTGTCAGAGGCTGTGCGAGCGGCGGCGGAAGAACTGCTGGCCTTCGCG	355
QY	101	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120



813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGCTTCTACGTCCAGTCCAGGGGATCCCGCAGGGTCTCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCAGCCTGTGCTACGCGCAGCATGGAGAACAACTGTTTGGCGGGATTTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTCGGCTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCAC 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGTGTGAATCTTCCCTGTAGAAGACGAGGCCCTGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACGCGCTATTTCCCTGTGTGGCGCTGTCTGTGTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACC 2932  
953 PheGlnSerValPheValAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTACACCGCGGCTTCAAGCGTGGAGGACATCGTCGCAACTCTTTGGGGTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCCTGTTCTTGGATTTCAGTGTGACAGCTCCAGAGCGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCCGTCGACGGCGTACAGGTTTCACGCAATGTGTGTGCGAGTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuLeuPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCATCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGTCTATCTCTGCACG 3172  
1033 AlaSerCysTyrValAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCTCATCTCTGAAAGCAAGAACGACGAGGATGTCTGGGGGCCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GGCGCGCGCGGCTCTGCGCTCCGAGGCGGTGACGTGGCTGTGTGCACCAAGCATCTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAAGCTGACTGCACACCGTGTCACTACGTGCCACTCTCTGGGTCTACTCAGCACACCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CACAGCAGCTGATGTGGAAGCTCTCCGGGGAGCAGCGCTGACTGCGCTGGAGCGCGACCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCACTGCGCTCAGACTTCAAGACCACTCTGGAC 3451

SULT 12

-09-128-354-1

Sequence 1, Application US/09128354

Patent No. 6337200

GENERAL INFORMATION:

APPLICANT: Morin, Gregg B.

APPLICANT: Geron Corporation

TITLE OF INVENTION: Human telomerase Catalytic Subunit Variants

FILE REFERENCE: 015389-003310US  
CURRENT APPLICATION NUMBER: US/09/128,354  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 08/851,843  
EARLIER FILING DATE: 1997-05-06  
EARLIER APPLICATION NUMBER: US 08/854,050  
EARLIER FILING DATE: 1997-05-09  
EARLIER APPLICATION NUMBER: US 08/911,312  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/912,951  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/915,503  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: US 08/974,549  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 08/974,584  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 09/052,864  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 4015  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (36)..(3454)  
OTHER INFORMATION: human telomerase reverse transcriptase (hTERT) cDNA  
US-09-128-354-1

Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13

US-09-042-460-2 (1-1122) x US-09-128-354-1 (1-4015)

Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCCGCGCGCTCCCGCTGCGAGCCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
Qy 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGGCCACGTTTGTGCGCGCCCTTGGGCCCCCAGGGCTGGCGGTGGTGAG 175  
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
Db 176 CGCGGGGACCGCGCGCTTCGCGCGCTGCTGCGCCAGTGTGCTGTGCTGCTGCCCTGG 235  
Qy 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GACGACG 295  
Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGGCCCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
Qy 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
Db 356 TTGCGGCTGTCTGACG 415  
Qy 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140  
Db 416 ACCTACCTGCGCCCAACACCGGTGACCGACCACTGTGGGGGAGCGGGGCTGGGGGCTGCTG 475



141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
146 CTGGCCCGCGTGGCGAGCGAGCTGGTTACCTGCTGGCAGCGCTGGCTTTGTG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla 180  
536 CTGTGGCTCCAGCTGCGCTTACAGGTGGCGGCGCGCTGTACAGCTCGCGCT 595  
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCATCAGGCCCGCGCCCGCCACACAGCTAGT---GGACCCGAGAGCGTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnLeuLysSerSerArgGlnGlnAlaPro 220  
650 -----TGGACGGCGCTGGAAACCATAGCTCAGGAGGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCGCTGGCTGCGACGCCCGGTGGAGAGCGCGGGGAGTGCAGCCGAAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGlnGlyPro--- 259  
749 CTGCGGTGGCCAGAGCCCGCGGTGGCGTGGCGTCCCTGAGCGGAGCGCGCGT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGCTCTGGCGCCACCCGGCGAGGCGGTGGACCGAGTGACCGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerLys 292  
869 GTGGGTGACCTGCCAGA-----CCCGCCGAGAGGCCACTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGCGCTCTCTGGCAGCGCCACTCCACCCCTCGTGGCGCGCCAGCACCGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCTCCACATCGCGGCCACCGTCCCTGGGACACCGCTTCTCCCGGTGTAGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTTACTCTCTCAGCGCAG---AAGGAGCAGCTCGCGCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTGAGGCCACAGCTGACTGGCGCTCGAGGCTCGTGGAGCCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCCCTGGATGCCAGGACTCCCGCAGGTGCCCCCTGCCCCAGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGCAATCGGCCCTGTTCTGGAGCTGCTTGGAAACACACGCGCAGTGGCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGGAGTGG-----GTCACCCCGAGCA 1321  
431 Leu-----Asn 432  
1322 GCGGGTGTGTGCGCGGAGAGCCCGAGGCTCTGTGGCGGCCCCCGAGGAGGAG 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCGCTGTGAGTGTGCTGCGCCAGCACAGCCCGCTGCGAGGTATC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGCGGGCTGCTGCTGCGCGGTGTGGCCCCCAGGCTCTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492

1502 AACGACCGCGCTTCTCAGGAACACCAAGAGTTCACTCCCTGGGAGCATGCCCCAAG 1561  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGCGAGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGCTTGGCTGGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCGAGGGTGGCTGTGTTCCGGCCGAGAGCACCGTCTGCGTGGAGAGATCTCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1682 AAGTTCCTGCTGCTGATGAGTGTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGlnSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrpSer 572  
1742 GTACGAGGAGCCAGCTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgLysLeuSer 592  
1802 AAGTTGCAAGCATTGGAAATCAGACAGCACTTGAAGAGGCTGAGCTGGCGAGGTCTCG 1861  
593 GlnGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGTCTGAGGAGCATCGGGAAGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCTATCCCAAGCTGACGGCTGCGCGGCTGTTGAACATGAGTCTGCTGCTGCGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 552  
1982 AGAACGTTCCGAGAGAAAGAGCGCGCGCTCTCACCTCAGGGGTGAAGCAGCTGTTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTCTCACTACGAGCGCGCGCGCGCGCTCTCTGGGCGCTCTGCTGCTGCTGCTGCT 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCAGAGGCTGCGCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGACTTGTCAAGGTGATGTGACGGCGCGCTACGACACCATCTCCCGAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTCATCGCGCATCATCAAA---CCCCAGAACACGTACTGCTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCAGAGCGCGCGCATGGGCGCGCTGCGGAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACTTTCAGACCTCCAGCGCTACACGACAGTTCGTGCTGCTGCTGCTGCTGCT 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGCTGCTCATGAGCAGAGCTCCTCCCTGAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGAGAGTGGCTCTTCGAGCGTCTCTACGCTTCTACGCTTCTACGCGCGCGCGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGCGCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852

2573 CTGCTCTGCACGCTGTCTACGGCGCATGGAGCAAGCTGTTTGGGGGATTGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGTCTCTCGCTTGGTGGATGATTTCTTGTGTGTGACACCTCACCTCACCCAC 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGlnTyrGlyCysMetIleAen 892  
2693 CGGAACCTTCTCAGACCTGCTCGAGGTGTCTCCTGAGTATGCTGCGGTGGAAC 2752  
893 LeuGlnIysThrValAlaAenPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGGGGAAGACAGTGTGAACTTCCCTGTGAAGACGAGGCCCTGGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGGCCACGGCTATTCCTCTGTGGCGCTGTCTGTGATACCCGAGCC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleIysThrSerLeuThr 952  
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGTCTCAC 2932  
953 PheGlnSerValPheLeuAlaHisGlyThrMetArgAsnIysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGCTTCAAGGCTGGAGAACATGCTGCGCAACTCTTTGGGGCTTGGGG 2992  
973 LeuIysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCTGTCTTCTGATTTGAGTTCAGGTGAGACAGCTCCAGAGGTGTGAC 3052  
993 AsnIleTyrIysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAGATCTCTCTGCTCAGCGGTACAGGTTTCACGATGTGTGTGACGCTC 3112  
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCAATTCATGAGAGTTTGAAGAACCCACATTTTCTGGCGTCACTCTGACAG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValIysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCTCATCTGAAAGCAAGAACGACGAGGATGCTGCTGGGGCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisIlePheLeuCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGGCTCTGCTCTCGAGCGCTGAGTGGCTGTGCGGCTCACTCAGGACAGCC 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGGTGTCACTAGTGTGCTGCTGCGGCTCACTCAGGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGAGCTGAGTGGAGCTCCCGGGAGCAGCTGCTGCTGAGCGCGCGGCGGCGG 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCACTGCCCTCAGACTTCAAGACCATCTGGAC 3451

## RESULT 13

S-09-675-321-1

Sequence 1, Application US/09675321

Patent No. 6440735

GENERAL INFORMATION:

APPLICANT: Gaeta, Federico C.A.

APPLICANT: Geron Corporation

TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune

TITLE OF INVENTION: Response to a Telomerase Antigen

FILE REFERENCE: 015389-003500PC

CURRENT APPLICATION NUMBER: US/09/675,321

CURRENT FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 60/112,006

PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: WO PCT/US99/06898

;; PRIOR FILING DATE: 1999-03-30  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 1  
;; LENGTH: 4015  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (56)..(3454)  
;; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)  
US-09-675-321-1

## Alignment Scores:

Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 4 Gaps: 13

US-09-042-460-2 (1-1122) x US-09-675-321-1 (1-4015)

Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGGTGCGCTCCCTGCTGCGAGCACCTACCGCGAG 115  
Qy 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTCTGCGCTGCGCACGTTCGTGCGCGCTGCGGCGCCCGAGGGCTGGCGGTGGTGAG 175  
Qy 41 ProGlyAspProIysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60  
Db 176 CGCGGGGACCGCGCGCTTTCGCGCGCTGCTGCGCGCTGCGCGCTGCTGCGCTGCGCTGG 235  
Qy 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GAGCGAGCG 295  
Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGCGCGAGTGTCTGAGAGGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
Qy 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
Db 356 TTGCGCTGTGCTGCG 415  
Qy 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaTrpMetLeuLeu 140  
Db 416 AGCTACTGCTCCACACGCTGACCGAGCGACTGCGGGGAGCGGGGCGTGGGGGCTGCTG 475  
Qy 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
Db 476 CTGCGCGCTGCG 535  
Qy 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
Db 536 CTGGTGGCTCCAGCTGCGCTACAGAGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCG 595  
Qy 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
Db 596 GCCACTCAGCG 649  
Qy 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGlnAlaPro 220  
Db 650 -----TGCAGACCGCGCTGGACCATAGCTCAGGAGCGCGCGCGCGCGCG 688  
Qy 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
Db 689 GTCCCCCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
Qy 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro---- 259

749 CTGCGTTGCCCCAAGAGGCCCCAGCGGTGGCGCTCCCTCGAGCGCGAGCGAGCGCCCGTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGGTCTCGGGCCCCACCGCGGAGAGCGGTGGACCGAGTACCGTGTCTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTCACCTGGCAGA-----CCCCCGAAGAGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GGTGCGCTCTGTGGCAGCGGCACTCCACCCATCCGTGGCGCGCAGCACACACGGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCATCGCGGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTACGGC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTCTACTCTCTCAGGCGAC---AAGGAGCAGCTGCGGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGCCCGCAGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCGCTTGATCGCAGGAGACTCCCGCAGGTGTCGCCCTGCCCGAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTGTGTTCTGAGCTGCTGGGAACACACGCGCAGTCCCGCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGACTGCGCGCTGCGAGCTGCG-----GTACCCCGAGCA 1321  
431 Leu-----Asn 432  
1322 GCGCGTGTCTGTGCCCGGAGAGCCCGCAGGCTCTGTGGCGGCGCCCGAGGAGGAGC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCTGTGTCAGTCTGCTCGCAGCAGCAGCAGCCCTCGCAGGTGATC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTTCGTGCGGCTGCTGCTGCGCGGCTGGTGGCCCCAGGSCCTCTGGGGCTCCAGGAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AACGACCGCGTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTCAGAGCTGACGTGGAGATGAGCTGCGGACTGCGCTTGGTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGGTGGTGTGTTCCCGCGCCAGACACCGCTCTGCGTGAGGAGATCCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGACTGGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGAGACACGCTTCAAAAGAACACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGGAATGACAGACGACTTTGAAGAGGGGTGCAGCTGCGGGAGCTGCTCG 1861

QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTCTAGGAGCATCGGAAGCAGCGCCCGCTCTGACTCCAGCTCCAGCTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCAATCCCAAGCTCTAGCGGCTCGCGCGCATGTGGAATGACTGCTGCTGGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGACGTTCCGACAGAAAGAGCGCGAGCGCTCTCACTCGAGGGTGAAGGCACCTGTTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGGTGCTCACTACAGCGGGCGCGCGCGCGCTCTGCGGCGCTCTGCTGTGGGC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGACGATATCACAGGCGCTGGCGCACCTTCTGCTGCTGCTGCGGCGCCAGGACCG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CCGCTGAGCTGTACTTGTCAAGGTGATGTACCGGCGCTGACACCATCCCCAG 2221  
QY 713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGGCTCAGGAGTCTCGCGCAGCATCATCAA---CCCCAGAACACAGTACTGCGTG 2278  
QY 733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
Db 2279 CGTGGTATCGCTGTGTCAGAAAGCGCCCGCTATGGCGACGTCCGCAAGCCCTCAAGAGC 2338  
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2339 CAGCTCTACCTTGACAGACCTCCAGCGCTATACGACAGTTCGTGGCTCACCTGCAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2399 GAGACC-----AGCCCGCTGAGGGATGCGCTGCTCATCGACAGAGCTCCCTCCCTGAAT 2452  
QY 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
Db 2453 GAGCGCAGAGTGGCTCTTCGACGCTTCTTACGCTTCTATGTGCCACACCGCTGGCGC 2512  
QY 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlnGlySerSerLeuSerThr 832  
Db 2513 ATCAGGGGCAAGTCTCTACGTCCTCAGTCCAGGGGATCCCGCAGGGCTTCCATCTCTCCAG 2572  
QY 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
Db 2573 CTGCTCTGACGCTGTCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGGCGG 2632  
QY 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
Db 2633 GACGGCTGCTCTCGCTTGTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2692  
QY 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
Db 2693 GCGAAACCTTCTCTCAGGACCTGTGTCGAGGTGTCCTGAGTATGCTGCTGCTGCTGCTGCTG 2752  
QY 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
Db 2753 TTGCGGAAGACAGTGGTGAATCTTCTCTGTAGAGAGAGCGCCCTCGGTGGCGGCTTTT 2812  
QY 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2813 GTTCAGATGCGCGGCCACCGCTATTCCTCTGTGGCGCTGTGCTGTGATATCCCGAGCC 2872  
QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2873 CTGAGGCTGACAGCGGACTACTCTCAGCTATGCCCGGACCTTCCATCAGAGCCAGTCTCACC 2932

953 PheInSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGAGGAAACATCGTCGCAAACTTTTGGGCTCTGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTCAACAGCTGTTCTTGGATTTCAGGTGAACAGGCTCCAGACGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCCGCTGCGAGGCGTACAGGTTTCACGCATGTGTGTCGAGTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGGGCTCATCTCTGACAGC 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCCATCTCTGAAGCCAAAGACGACGGGATGTCGTGGGGCCCAAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisIleTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GGGCCCGCGGCTCTGCTCCCTCGAGGCGGTGCACTGCTGTGCACCAACGATTCCTG 3252  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1099  
3293 CTCAGCTGACTGCACACCGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGAGGAGCTGAGTGGAGAGCTCCCGGGGACGACGCTGCTGCTGCTGCTGCTGCTGCT 3412  
1110 AspProAlaLeuSerThrAppPheGlnThrIleLeuAsp 1122  
3413 AACCGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3451

## RESULT 14

-09-052-919-1

Sequence 1, Application US/09052919

Patent No. 6444650

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Antisense Compositions for Detecting and

Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/052,919

FILING DATE: 31-MAR-1998

CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 23-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "human telomerase reverse  
transcriptase (hTRT)"  
US-09-052-919-1

Alignment Scores:  
Seq. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conserved: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13

US-09-042-460-2 (1-1122) x US-09-052-919-1 (1-4015)

Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCCGGCGCTCCCGCTGCCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
Qy 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCCGCTGCCACCGTTCGTGCGGCGCTGCGGCGCCCGAGGCTGCGCTGTCAG 175  
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGGACCCCGCGGCTTTCCGCGCGCTGTGCGCCAGTCCCTGCTGTGCGGCTCGG 235

1271 GGGGTGCTCTCTCAAGACGCACTCCCGCTGCAGCTGG- - - - -GTCACCCCAGCA 1321

431 Leu- - - - -Asn 432

1322 GCCGTGTCTGTGCCGGGAGAGCCCGAGGGCTCTGTGGCGGCCCGCCAGGAGGAGC 1381

433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrioGlnValTyr 452

1382 ACAGACCCCGCTGCTGCTGGTGCAGCTGCTCCGCGACACAGACAGCCCTTGCAGGTGTAC 1441

453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrioGlyThrArgHis 472

1442 GGCCTCGTGGCGGCTGCTGCGCGGCTGGTGGCCCAAGGCTCTGGGGCTCCAGGCAC 1501

473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheLeuSerLeuGlyLysTyrGlyLys 492

1502 AACGAACCCGCTCTCTCAGGAACCAAGAAAGTTTCATCTCCCTCGGGAAGCATGCCAAG 1561

493 LeuSerLeuGlnGluLeuMetTrioLysMetLysValGluAspCysHisTrioLeuArgSer 511

1562 CTCTCGCTGCAGGAGCTCGAGTGGAGATGAGCGTGGCGGACTGCGCTGGCTGGCGCAGG 1621

513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532

1622 AGCCAGGGGTGGCTGTGTTCGGCGCGACAGACCGCTCGGTGAGGAGATCTCGGCC 1681

533 ThrPheLeuPheTrioLeuMetAspThrTyrValValGlnLeuArgSerPhePheTyr 552

1682 AAGTCTCTGCACTGGCTGATCAGTGTACGTGCTCGAGCTGCTCAGGTCTTCTTTTAT 1741

553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrioSer 572

1742 GTCACGAGACCCAGTTCACAAAGAACAGSCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801

573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592

1802 AAGTGTCAAGCATTTGAAATCAGACAGCACTTGAAGAGGTGCAGCTCGGAGAGTGTGC 1861

593 GlnGluGluValArgHisGlnAspThrTrioLeuAlaMetProIleCysArgLeuArg 612

1862 GAACAGAGGTTCAGCAGCATCGGAGACCGAGCGCCCTGCTGACGTCCAGATCCGCG 1921

613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632

1922 TTCAATCCCAAGCCTGACGGCTCGCGCCGANTGTAAATGCAATGACTAGCTGTGGAGACC 1981

633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysPheLeuPhe 652

1982 AGAAGTTCGCGAGAAAGAGAGCGCGCGAGCGTCTCACTTCAGAGGTGAAGGCATGTTC 2041

653 SerMetLeuAsnTyrGluArgThrLysPheHisProHisLeuMetGlySerSerValLeuGly 672

2042 AGCGTGTCAACTACGAGCGGCGCGCGCCCGCGCTCTCGGGCGCTCTGTGTGGCG 2101

673 MetAsnAspIleTyrArgThrTrioArgAlaPheValLeuArgValArgAlaLeuAspGln 692

2102 CTGAGCATATCCACAGGGCTCGCGCACCTCGTCTGCGTGTGGGCGCGAGACCCG 2161

693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712

2162 CGCCTTGAGCTGACTTGTCAAGTGTGATGTACGGGCGGTACGACCATCCCCCAG 2221

713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732

2222 GACAGGCTCAGGAGGTTCATCGCAGCATCAAAA- - - - -CCCCAAGACGCTACTGCGTG 2278

733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752

2279 GGTGGTATGCGTGGTTCAGAGGCGGCCCATGGGACGTCCGCAAGGCCTTCAAGAC 2338

753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```



S-08-912-951-1

Ligment Scores:

Red. No.:	0	Length:	4015
Score:	3505.00	Matches:	719
Percent Similarity:	72.94%	Conservative:	122
Est Local Similarity:	62.36%	Mismatches:	260
Very Match:	59.40%	Indels:	52
B:	4	Gaps:	13

S-09-042-460-2 (1-1122) x US-08-912-951-1 (1-4015)

```

Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
b 56 ArgCGCGCGCTCCCGCTGCGAGCGTGGCTGCTGCTGCTGCGACCACTACCGCGAG 115
Y 21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
b 116 GTGCTGCGCTGGCCACAGTTCGTGCGCGCTGGCGGCCCGCCAGCGCTGGCGCTGCGAG 175
Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTTP 60
b 176 CCGCGGAGACCGCGCGCTTCGCGCGCTGTGTGCGCCAGTGCCTGTGTGCGCTGCG 235
Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
b 236 GACGACGCGCGCGCGCGCGCTTCCTTCGCGCAGGTGTCTGCTGCTGAAGGAGCTG 295
Y 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
b 296 GTGGCGCGAGTGTCTCAGAGGTGTGCGAGCGCGCGCGAGAACGTGCTGGCTTCGCG 355
Y 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
b 356 TTGCGCTGCTGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCTTCACACAGCGTGGC 415
Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTTPMetLeuLeu 140
b 416 AGCTACCTGCTCCCAACACAGGTGACCGACCACTGGCGGGGAGCGGGCGTGGCTGCTG 475
Y 141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
b 476 CTGCGCGCTGGCGGACAGCTGTGTGTACCTGTGGCGCGCGCTGCGCGCTGCTTTGTG 535
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
b 536 CTGTGTGCTCCCGAGCTGCGCTTACAGGTGTGCGCGCGCGCGCTGTACAGCTCGCGCT 595
Y 181 ThrThrAspIleTTPProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
b 596 GCGACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGCGCTGGGA 649
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
b 650 -----TGCGACGGGCTGGAACCATAGCGTCAGGAGCGCGCGCGCGCG 688
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
b 689 GTCCCCCTGGGCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAAGT 748
Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProAlaGluGluGlyPro--- 259
b 749 CTGCGCTTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTTP--- 272
b 809 GGGCAGGGTTCCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
b 869 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGGCGCGCGCGCGCGCGCGCG 913
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysPheProSer 311

```

```

Db 914 GGTGCGCTCTCTGCGACGCGCCACTCCACCCATCGTGGCGCGCGACACACGCGGCG 973
Qy 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---Pheile 330
Db 974 CCCCATTCCACATCGCGGCGCACCATGCTCCGAGACACAGCTTGTCCCGCGGTGTACGCC 1033
Qy 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACCTCTCTACTCTCTCAGCGAC---AAGAGCAGCTGCGCGCTCTCTTC 1090
Qy 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeuPhe 370
Db 1091 CTACTCAGCTCTCTGAGGCCACGCTGCTGGCGCTCGGAGCTCGTGAGACCACTCTT 1150
Qy 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1151 CTGGTTTCCAGGCGCTGGATGCCAGGACTTCCCGCAGGTTCGCCCGCTGCCCGCGCGC 1210
Qy 391 TyrTTPGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
Db 1211 TACTGGCAATGGCGCGCTGTTCTGGAGCTGCTTGGGAACACACGCGCAGTGGCCCTAC 1270
Qy 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
Db 1271 GGGGTGCTCTCAAGACGCACTCCCGCTCGAGCTGCG-----GTACCCCGACGA 1321
Qy 431 Leu-----Asn 432
Db 1322 GCCGTGTCTGTCGGGAGAACCCAGGGCTCTGTGCGCGCGCGCGCGAGGAGGAC 1381
Qy 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTTPGlnValTyr 452
Db 1382 ACAGACCCCGCTCGCTGTGTGCTGCTCGCGACGACAGCAGCGCGCTGGCAGGTGTAC 1441
Qy 453 GlyPheLeuGluAlaCysLeuCysLysValValSerAlaSerLeuTTPGlyThrArgHis 472
Db 1442 GCGTTCGTGGCGCGCTCGCTGCGCGCTGTGTGCCCGCGCGCGCTGTGGGCGCTCCAGGCAC 1501
Qy 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492
Db 1502 AACGAACGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAG 1561
Qy 493 LeuSerLeuGlnGluLeuMetTTPLysMetLysValGluAspCysHisTTPLeuArgSer 512
Db 1562 CTCTCGCTGAGGAGTGTGCTGGAAGATGAGCGTGGGAGCTGCGCTGGCTGGCGCAG 1621
Qy 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
Db 1622 AGCCAGGGGTGGCTGTGTTCGGCGCGCAGACACCGCTGCGTGGAGAGATCTCTGGCC 1681
Qy 533 ThrPheLeuPheTTPLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
Db 1682 AAGTTCCTGCACTGGCTGATGAGTGTGTGCTGCTGAGCTGCTCAGGTCTTTCTTTAT 1741
Qy 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTTPSer 572
Db 1742 GTCACGAGACCACTTCAAGAACAGAGCTCTTTTCTACCGGAGAGTGTCTGGAGC 1801
Qy 573 LysLeuGlnSerIleGlyValArgGlnHisLeuLeuArgValArgLeuArgGluLeuSer 592
Db 1802 AAGTTGCAAGCAATGGAATCAGACAGCACTTGAAGAGGTGCGAGCTGCGGAGCTGCG 1861
Qy 593 GlnGluGluValArgHisHisGlnAspThrTTPLeuAlaMetProIleCysArgLeuArg 612
Db 1862 GAAGCAGAGGTGAGCGAGCTCGGGAAGCGCGCGCGCTGTGAGCTCCAGCTCCCGC 1921
Qy 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
Db 1922 TTCATCCCAAGCTGAGCGCTCGCGCGCATGTTGAACTGAGCTACGTGCGTGGAGGC 1981
Qy 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652

```

1982 AGAAGCTTCGACAGAAAAGAGGGCCGAGCGCTCTCACCTCGAGGGTCAAGGCACTGTC 2041  
553 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACAGAGCGGGCGGCCCGCCGCTCTCTGGCGCCTCTGTGTGGGC 2101  
673 MetAsnAspIleTyrArgThrTTPArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGGCTCGGCACCTCTGTGTCTGTGTGGGGCCCGAGACCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTTCACAGGTGGATGTACGGGCGGTACGACACCATCCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCACGGAGTTCATCGCCAGCATCATCAA- -CCCCAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCCGCTGTCTCAAGAGCGGCCCATGGGCACGTCCGCAAGGCCCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACCTTGACAGACTCCAGCGGTATACGCGACAGTTGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC- - - - -AGCCCGCTGAGGGATCGCTCGTCAACGACAGAGTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
2453 GAGCCAGAGTGGCTCTTCGACGTCTTCCTACGCTTCATGTGCCACACCGCGTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTACGTCCAGTCCAGGGGATCCCGACAGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCACCTGTGTCTACGGCAGATGAGAACCAAGCTTTTGGCGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872  
2633 GACGGGTGTCTCTGCTGTGTGGTGGATGATTTCTGTGTGACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCTCAGACCCCTGGTCCGAGGTGTCTCTGAGTATGGCTGTGTGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyValAlaPro 912  
2753 TTGCGAAGACAGTGGTGAACCTTCCCTGTAGAGACAGAGGCCCTGGGTGGACGGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGGCCACCGCCCTATTCCCTCGTGGCGCTGTGTGTGATACCCGAC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGACAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCAC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAGGCTGGAGGACATGCGTCGCAACTCTTTGGGGTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACACGCTGTTTCTGGATTTGAGGTGAACAGAGCTCCAGACGGGTGTGCAC 3052  
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCTGCGGGGTACAGGTTTACGCGATGTGTGTGCGCTC 3112

QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCAATTTCATCACAAGTTTGAAGAACCCACATTTTCTGCGGTCATCTCTGACACG 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCTCTGTACTCCATCTGAAGCCAAAGAACGACGAGGATGTGCTGGGGGCCAAG 3232  
QY 1053 GlySer- - - - -PheProProGluAlaAlaHisIlePheLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GGGCGCGCGGCCCTCTGCCCTCCGAGGCCGTGAGTGGCTGTGCCACCAAGCATTTCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAGCTGACTCGACACCGTGTACCTACGTGCCACTCTCTGGGGTCACTCAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGAGCTGAGTCGAAGCTCCCGGGGACGAGTGTGCTGCCCTGGAGGCCGACGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGCACTGCCCTCAGACTTCAAGACCATCCTGGAC 3451

Search completed: February 1, 2004, 19:35:25  
Job time : 360 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

n on: February 1, 2004, 18:05:20 ; Search time 877 Seconds  
(without alignments)  
4662.705 Million cell updates/sec

tle:  
xrefct score: 5901  
quence: 1 MTRAPRCPAVRSILRSRYRE.....TILKAADPALSTDFQITLD 1122

oring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

arched:  
2434939 seqs, 1822278265 residues

tal number of hits satisfying chosen parameters: 4869878

nimum DB seq length: 0  
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

mmmand line parameters:  
MODEL=frame+p2n.model -DEV=xlh  
B=/cgn2\_1/USPTO.spool/US09042460/runat\_01022004\_150553\_17127/app\_query.fasta\_1.1287  
B=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
COPI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
RANS=human40.cdd -LITS=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
HR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0  
AXLEN=200000000 -USER=US09042460@cgn 1.1 586 @runat\_01022004\_150553\_17127  
ICPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
JONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

itabase : Published Applications NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Query	Match	Length	ID	Description
-------	-------	--------	----	-------------

1	3505	59.4	3396	10	US-09-749-728B-32	Sequence 32, Appli
2	3505	59.4	4015	9	US-09-733-294A-3	Sequence 3, Appli
3	3505	59.4	4015	10	US-09-990-080-1	Sequence 1, Appli
4	3505	59.4	4015	10	US-09-843-676-224	Sequence 224, App
5	3505	59.4	4015	10	US-09-953-052-1	Sequence 1, Appli
6	3505	59.4	4015	15	US-10-053-758-224	Sequence 224, App
7	3505	59.4	4015	15	US-10-208-243-1	Sequence 1, Appli
8	3505	59.4	4015	15	US-10-054-295-224	Sequence 224, App
9	3505	59.4	4015	15	US-10-054-611-224	Sequence 224, App
10	3505	59.4	4015	15	US-10-105-963-1	Sequence 1, Appli
11	3505	59.4	4015	15	US-10-044-692-1	Sequence 1, Appli
12	3505	59.4	4015	15	US-10-044-692-1	Sequence 1, Appli
13	3498	59.3	4027	12	US-10-385-882-1	Sequence 1, Appli
14	3496	59.2	13766	13	US-10-105-616-1	Sequence 1, Appli
15	3496	59.2	13766	13	US-10-105-616-1	Sequence 6, Appli
16	3473	58.9	8742	13	US-10-105-616-6	Sequence 173, App
17	3287	55.7	4029	11	US-09-843-676-173	Sequence 173, App
18	3287	55.7	4029	11	US-09-438-486-173	Sequence 173, App
19	3287	55.7	4029	15	US-10-053-758-173	Sequence 173, App
20	3287	55.7	4029	15	US-10-054-295-173	Sequence 173, App
21	3287	55.7	4029	15	US-10-054-611-173	Sequence 173, App
22	3276.5	55.5	3855	15	US-10-044-692-4	Sequence 4, Appli
23	3276.5	55.5	3855	15	US-10-044-692-4	Sequence 11, Appli
24	2132	36.1	1866	15	US-10-294-778-11	Sequence 3, Appli
25	1844.5	31.3	2176	15	US-10-044-692-3	Sequence 3, Appli
26	1844.5	31.3	2176	15	US-10-044-692-3	Sequence 100, App
27	1786.5	30.3	2171	10	US-09-843-676-100	Sequence 100, App
28	1786.5	30.3	2171	10	US-09-766-253-100	Sequence 100, App
29	1786.5	30.3	2171	11	US-09-438-486-100	Sequence 100, App
30	1786.5	30.3	2171	15	US-10-053-758-100	Sequence 100, App
31	1786.5	30.3	2171	15	US-10-054-295-100	Sequence 100, App
32	1786.5	30.3	2171	15	US-10-054-611-100	Sequence 100, App
33	1511.5	25.6	1314	15	US-10-294-778-9	Sequence 9, Appli
34	1497	25.4	1311	15	US-10-294-778-1	Sequence 1, Appli
35	1336	22.6	4200	15	US-10-044-692-6	Sequence 6, Appli
36	1336	22.6	4200	15	US-10-044-692-6	Sequence 1, Appli
37	1336	22.6	15418	9	US-09-783-203-1	Sequence 1, Appli
38	1336	22.6	15418	10	US-09-934-427A-1	Sequence 1, Appli
39	1336	22.6	15418	11	US-09-935-419A-1	Sequence 1, Appli
40	1336	22.6	15418	15	US-10-141-220-1	Sequence 1, Appli
41	1336	22.6	15418	15	US-10-023-969-1	Sequence 1, Appli
42	1336	22.6	15418	15	US-10-206-447-1	Sequence 1, Appli
43	1336	22.6	51552	9	US-09-733-294A-30	Sequence 30, Appli
44	972.5	16.5	873	13	US-10-282-960-1	Sequence 1, Appli
45	702	11.9	519	13	US-10-282-960-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-09-749-728B-32  
; Sequence 32, Application US/09749728B  
; Patent No. US20020142457A1  
; GENERAL INFORMATION:  
; APPLICANT: Umezawa, Akihiro  
; APPLICANT: Hata, Jun-ichi  
; APPLICANT: Fukuda, Keiichi  
; APPLICANT: Ogawa, Satoshi  
; APPLICANT: Sakurada, Kazuhiro  
; APPLICANT: Gojo, Satoshi  
; APPLICANT: Yamada, Yoji  
; TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIO  
; FILE REFERENCE: 00766.000043  
; CURRENT APPLICATION NUMBER: US/09/749,728B  
; CURRENT FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: H11-372826  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: PCT-JP00-01148  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: PCT-JP00-07741  
; PRIOR FILING DATE: 2000-11-02  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: Patentin Ver.2.0



```
613 PheileProLysProAnglyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1867 TTATCCCAAGCTGAGCGGCTGCGGCCGATTTGTGAACATACGTCGTGGGAGCC 1926
633 ArgAlaLeuGlyArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1927 AGAAGCTTCCGACAGAAAGAGAGGCGGAGCGTCTACCTCGAGGTGAAGCAGCTGTC 1986
653 SerMetLeuAenTyrGluArgThrLysHisProHisLeuMetGlySerValLeuGly 672
1987 AGCGTGCTCAACTACGAGCGGCGCGCGCGCCGCTCTCTGCGCGCTCTGTGTGGGC 2046
673 MetAenAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2047 CTGACGATATCCACAGGCGCTGCGGACCTTGTGCTGCTGTGCGGCCAGGACCG 2106
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2107 CCGCTGAGCTGTAATTGTCAAGGTGATGTGACGGGCGCGTACGACACCATCCGCCAG 2166
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2167 GACAGGCTCAGGAGGTATCGCCAGCATCATCAAA---CCCGAACACGCTACTCGGTG 2223
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2224 CGTCGGTATGCGGTGTCAGAGGCGGCCCATGGGCACTGCGCAAGGCTTCAAGAGC 2283
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2284 CACGTCTCTACTCTTGACAGACCTCCAGCCGCTACATGCGACAGTTCGTGGCTCACCTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792
2344 GAGACC-----AGCCCGCTGAGGAGTGGCGTGTATCGAGCAGAGCTCCCTCGAT 2397
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812
2398 GAGGCGACAGTGGCGCTTCTGAGCGTCTTCTACGCTTCATGTGCCACACCGCGTGGCG 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2458 ATCAGGCGCAGTCTCTAGCTCAGTGGCAGGAGATCCCGAGGCTCCATCTCTCCAG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2518 CTGCTCTCAGCGCTGTGTACGGCGACATGAGAGCAAGCTGTTTGGCGGATTCGGCGG 2577
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2578 GACGGGCTGCTCTCTGCGGTTTGGTGGATTTCTTGTGTGATGCTCCTCACCTCACCCAC 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2638 GCGAAACCTTCTCAGGACCTGTCGAGGTTGCTCGAGTATGCTGCTGCTGCTGCTGCTG 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912
2698 TTGCGGAGACAGTGGTGAATCTCCCTGTAGAAGACGAGGCGCTGCTGCTGCTGCTGCT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2758 GTTCAGATGCGGCGCCAGCGCTATTCTCCCTGCTGCGGCTGCTGCTGCTGCTGCTGCTG 2817
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2818 CTGAGGTGCGAGAGCGACTACTCCAGCTATGCCCGACCTCCATCAGAGCGAGCTCTCAC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2878 TTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAACTCTTTTGGGCTTTGGCG 2937
```

```
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
Db 2938 CTGAGGTGCACAGCTGTTTCTGAGTTTGCAGGTGAACAGCTCCAGCGTGTGCACC 2997
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
Db 2998 AACATCTACAAGATCTCTCTGCTGCGAGGGGTACAGTTTCAAGCATGTGTGTGTCAGCTC 3057
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3058 CCATTTTCATCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGCTCATCTCTGACAGC 3117
QY 1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
Db 3118 GCCTCCTCTGCTACTCCATCTGTAAGCCCAAGACGAGGATGCTCGTGGGCGCAAG 3177
QY 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
Db 3178 GCGCGCGCGCGCTCTGCTGCGAGGCGGTGCAGTGGCTGTGCCACCAAGCATTTCTCGT 3237
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
Db 3238 CTACAGCTGACTCGACACCGTGTACCTACCTAGTGCACCTCTGGGCTCACTCAGGACGCC 3297
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
Db 3298 CAGACGCGAGCTGAGTGGAGGCTCCCGGGGAGCGAGCTGACTGCGCTGGAGGCGCGAGCC 3357
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3358 AACCGCGCTGCTCCCTCAGACTTCAGAGCATCTCTGGAC 3396

RESULT 2
US-09-733-294A-3
; Sequence 3, Application US/09733294A
; Patent No. US20020045588A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William Gaarde
; APPLICANT: Susan M. Freier
; APPLICANT: Edward V. Wanciewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0527
; CURRENT APPLICATION NUMBER: US/09/733,294A
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/572,423
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 108
; SEQ ID NO 3
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)....(3454)
US-09-733-294A-3

Alignment Scores:
Pred. No.: 0 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
Query Match: 59.40% Indels: 52
DB: 13

US-09-042-460-2 (1-1122) x US-09-733-294A-3 (1-4015)
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATCCGCGCGCTCCCGCTGCGAGCGGTGCGCTCCCTGCTGCGGAGCCACTACCGCGAG 115
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgLeuValGln 40
```

b 116 GTGTCGCGCTGGCCACGCTTGTGCGCGCGCTGGGCCCCCAGGGCTGGCGCTGGTGACAG 175  
y 41 ProGlyAspProLysIleTyArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
b 176 CGCGGGACCCGCGCGCTTCCGCGCGCTGGTGGCCACGTCGCTGTGTGCGTCCCTGG 235  
y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80  
b 236 GAGCAGCGCGCGCGCGCGCGCTTCTTCCGCGCAGGTGCTTCCCTCCAGAGAGCTG 235  
y 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
b 296 GTGGCCCGAGTGTCCAGAGGCTGTGCGAGCGCGCGCGAAGACGTCGTGGCTTCGGC 355  
y 101 PheGluLeuLeuAsnGluAlaArgGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTGCGCGCTGTGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 415  
y 121 SerTyLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACTGCCCAACACGCTGACCGACGACTGCGGGGAGCGGGGCTGCTG 475  
y 141 LeuSerArgValGlyAspLeuLeuValTyLeuLeuAlaHisCysAlaLeuTyLeu 160  
b 476 CTGCGCGCGTGGCGACACGTCGTGTTCACCTGCTGCGACGCTGCGCGCTTTTGTG 535  
y 161 LeuValProProSerCysAlaTyGlnValCysGlySerProLeuTyGlnIleCysAla 180  
b 536 CTGCGTGTCCACGTCGCCCTACGAGTGTGCGGCGCGCGCTGTACCGACTCGGCGCT 595  
y 181 ThrThrAspIleTrpProSerValSerAlaSerTyArgProThrArgProValGlyArg 200  
b 596 GCCACTCAGCGCGCGCGCGCGCGCGCTAGT---GGACCCCGAAGCGTCTGGGA--- 649  
y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TGCGACCGGGCTCGAACCATTAGCTGCGGAGCGCGCGG 688  
y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerSer 240  
b 689 GTCCCCCTGGCGCTGCCAGCCCGCGTGGAGAGCGCGCGCGCGCGCGCGCGCGAGT 748  
y 241 ValProSerAlaLysLysAlaArgCysTyProValProArgValGluGluGlyPro--- 259  
b 749 CTGCGTGTCCCAAGAGCGCCAGCGTGGCGTGCCTGCGCGCGCGCGCGCGCGCGCTT 808  
y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
b 809 GGCAGGGGTCTGGGCCACCCGCGCAGACGCTGGACCGAGTGCCTGTTCTGT 868  
y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 913  
y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GTGGCGCTCTTGGCAGCGCCACTCCACCCATCCGTGGCGCGCGCGCGCGCGCGCGG 973  
y 312 SerThrSerLeuLeuSerProProArgGlnInAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCGCATCCACATCGCGGCGCACCGCTGCTGGACACGCTTGTCCCGCGTGTAGCC 1033  
y 331 GluThrArgHisPheLeuTySerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACCTTCTCTACTCTCTCAGCGAGAC---AAGAGCAGCTCGCGCGCTTTC 1090  
y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTCTGAGCGCCAGCCTGACTGCGCTCGGAGCTCGTGGAGACCATCTTT 1150  
y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
b 1151 CTGGGTTCCAGCGCGCTGATGCCAGGAGTTCGCCCGCGAGTTCGCCCGCTGCCCGCGC 1210

QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTy 410  
Db 1211 TACTGGCAATATGGGCGCTTCTTCTGGAGCTCTTGGGAACACCGCAGTGGCCCTTAC 1270  
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
Db 1271 GGGTGTCTCTCAAGACGCACTGCCGCTGCGAGTGGC-----GTACCCCGCAGCA 1321  
QY 431 Leu-----Asn 432  
Db 1322 CCCGCTGTCTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAC 1381  
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTy 452  
Db 1382 ACAACCCCGCTGCTGTGCGAGCTGCTCCGCCAGCAGCAGCAGCCCTGGCAGGTGAC 1441  
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTTCGTGCGGGCTGCTGCGCGCGCTGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGC 1501  
QY 473 AsnGluArgArgPhePheLeuLeuLysPheIleSerLeuGlyLysTyGlyLys 492  
Db 1502 AACGACCCCGCTTCTCCAGAACACCGAAGATTTATCTCCCTGGGAGAGCATGCCAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTGCGAGGCTGACGTGGAAGTGGCGTGGAGTCTGCTGCTGCTGCTGCTGCTG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCGAGGGTGGCTGTGCTGCGCGCGCAGAGACCGCTCTGCGTGAGAGATCTCTGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyValValGlnLeuLeuArgSerPhePheTy 552  
Db 1682 AAGTTCCTGCACTGCTGATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyArgLysSerValTrpSer 572  
Db 1742 CTCACGGAGACACGCTTTCAAAGAACAGAGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AGTTGCAAGCANTGGATTCAGACACACTTGAAGAGGGTGCAGCTGCGGAGCTGTG 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTGAGCAGCATCGGGAAGCCAGCGCCCGCTGCTGAGCTCCAGACTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTySerMetGlyThr 632  
Db 1922 TTCATCCCAAGCTTGACGGCTGCGCGCGCGCGCGCTGTAACATGACTACGTCTGGAGGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAACGTTCCGACAGAAAAGAGCGCGCGCTCTACCTCGAGGGTGAAGGACCTGTTTC 2041  
QY 653 SerMetLeuAsnTyArgLeuArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCAACTACGACGCGCGCGCGCGCGCGCTCTCTGCGCGCTCTGCTGCTGCGG 2101  
QY 673 MetAsnAspIleTyArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGAGCATATCCACAGGGCTTGGCGCACCTTCTGCTGCTGCTGCTGCGCGCGCGCGC 2161  
QY 693 ThrProArgMetTyPheValLysAlaAspValThrGlyAlaTyArgAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGTACTTGTCAAGTGTGATGTCGCGCGCGCTACGACACCATCCCGCAG 2221  
QY 713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyCysIle 732  
Db 2222 GACAGGCTCAGCGAGGTCTATCGCCAGCATCATCAAA-----CCCCAGAACAGTACTGGGTG 2278





Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGCTGGCTCCAGCTGCGCTACAGGTGCGGCGCGCTGTACAGCTCGCGCT 595  
Y 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCACCTCAGCG 649  
Y 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGlnAlaPro 220  
b 650 -----TGCGAACGGCGCTGGAACCATAGCTGAGGAGCGCGG 688  
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
b 689 GTCCCGCTCGGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
Y 241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTGCGCTTCCCGAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
Y 260 -----HisArgGlnValLeuProThrProSerArgLysSerTrp--- 272  
b 809 GGGCAGGGCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGCGCACCTCTTTGGAG 913  
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GGTGGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
Y 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033  
Y 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGCAC---AAGGAGCAGCTGCGCGCTCTC 1090  
Y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
Y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
b 1151 CTGGGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210  
Y 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
b 1211 TACTGGCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270  
Y 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
b 1271 GGGGTCTCTCAAGACGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1321  
Y 431 Leu-----Asn 432  
b 1322 GCGGTGTCTGTGCGCGGAGAACCCAGCGCTCTGTGGCGCGCGCGCGCGCG 1381  
Y 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
b 1382 ACAGACCGCGCTGGCTGGTGGAGTGTCTCCGACGACACAGCGCGCGCGCG 1441  
Y 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
b 1442 GGCTTGGTGGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501  
Y 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
b 1502 AACGAACGCGCGCTTCTCAGGAACACCAAGAGGTTCATCTCCCTGGGAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
DB 1562 CTCTCGCTCAGAGAGCTGACGTGAAGATGACGCTGCGGAGCTGCTTGGCTGCGCAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgIleLeuAla 532  
DB 1622 AGCCAGGGGTGGCTGTGTTCCGCGCGCAGACACCGTCTCGTGAGGAGATCTCGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
DB 1682 AAGTTCCTGACCTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGTCTTCTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
DB 1742 GTACGAGACACAGTTCAAAAGAACAGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
DB 1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGACAGCTGCGGAGCTGTGC 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
DB 1862 GAACAGAGGTGAGCAGCATCGGAAGCGCGCGCGCTCTGAGCTCCAGACTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
DB 1922 TTATCCCCAAGCTTACCGGGCTGCGCGCGCTTGTGAACATGAGCTACGTCTGGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
DB 1982 AGAAGTTCGCGAGAGAAAGAGCGCGCGAGCTCTCCTCGAGGGTGAAGCACTGTTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
DB 2042 AGCGTGTCTCAACTACGAGCGCGCGCGCGCGCGCTCTCTGGCGCGCTCTGTGTGGGC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
DB 2102 CTGACCATATCCAGAGGCGCTGGCGACCTCTGCTGCTGCTGCGGGCCCGAGACCGG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValArgAspAlaIleProGln 712  
DB 2162 CCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGCTACGACACCATCCCGCAG 2221  
QY 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
DB 2222 GACAGGCTCAGGAGGTATCCGCGCATCATCAAA---CCCGAGAACACGACTACTCGGTG 2278  
QY 733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
DB 2279 CGTGGTATGCGGTGTCTCAGAGCGCGCGCGCGCGCGCTCCGCAAGCCCTTCAAGAGC 2338  
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
DB 2339 CACGTCTCTACTTACAGACCTCCAGCGCTACATCGACAGCTCTGCTGCTCCTCCCTGAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
DB 2399 GAGACC-----AGCCCGCTGAGGATGCGCTCATCGAGCGAGCTCTCCCTCTGAAT 2452  
QY 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
DB 2453 GAGCGCAGCAGTGGCTCTTCGACGTCTCTACGCTCTTATGTGCCACCGCGCTGCGC 2512  
QY 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
DB 2513 ATCAGGGCAAGTCTACGTCCAGTCCAGGGATCCCGAGGGCTCCATCTCTCCAGC 2572  
QY 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
DB 2573 CTGCTCTGACGCTGTCTACGCGGACATGGAGACAAGCTGTTCGGGGATTCGGCGG 2632  
QY 853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872

||||| 2633 GACGGGCTGCTCTGCGGTTGGTGGAGATTTCTTGGTGACACCTCACCTCACCCAC 2692  
||||| 873 AlAlysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
||||| 2693 GCGAAACCTTCTCTCAGGACCCCTGGTCCGAGGTGTCTCCTGAGTATGGTGTGGTGAAC 2752  
||||| 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
||||| 2753 TTGGGAGAGACAGTGGTGAATCTTCCCTGTGTAAGACAGAGCCCTGGTGGCACGGCTTTT 2812  
||||| 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
||||| 2813 GTTCAGATGCGGGCCACCGCTATTCCCTGCTGGTGGCTGCTGCTGATACCCGGACC 2872  
||||| 933 LeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
||||| 2873 CTGAGGTGCGAGCGACTACTCCAGTATGCCCGGACCTCCATCAGACCCAGTCTCACC 2932  
||||| 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
||||| 2933 TTCAACCGCGCTTCAAGGCTGGGAGGAACATCGTGCAGAACTCTTTGGGGTCTTGGG 2992  
||||| 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
||||| 2993 CTGAAGTGTACAGCCCTGTTCTGGATTTCAGGTGACAGCCCTCAGACGGTGTGCACC 3052  
||||| 993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
||||| 3053 AACATCTACAGATCTCTCTGCTGAGCGGTACAGTTTCAGCATGTGTGTGAGCTC 3112  
||||| 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
||||| 3113 CCATTTCATCAGCAAGTTTGAGAACCCACATTTTCTGCGCGCTCATCTCTCACAG 3172  
||||| 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
||||| 3173 GCTTCCTCTCTACTCATCTCTGAAAGCCAGAACGAGGAGTGTCTGGGGGCCAAG 3232  
||||| 1053 GlySer-----PheProProGluAlaAlaHisTyrPheCysTyrGlnAlaPheLeu 1069  
||||| 3233 GCGCGCGCGCGCTCTGCTGCTGCGAGCGGTGAGTGGTGTGCTGCGCCAGCAAGCATTCCTG 3292  
||||| 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
||||| 3293 CTCAAGCTGTACTGCACACCGGTGTACCTAGTGCCTCTCTGGGGTCACTCAGGACAGCC 3352  
||||| 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
||||| 3353 CAGACGCGAGCTGAGTGGAGCTCCCGGGGACGAGCTGACTGCTGCTGGAGCGCGAGCC 3412  
||||| 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
||||| 3413 AACCGGCACTGCGCTCAGATCTCAGACCATCTCAGACCATCTCTGCAC 3451

## RESULT 4

S-09-843-676-224

Sequence 224, Application US/09843676

Patent No. US20020164786A1

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. US20020164786A1e1 Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/843,676  
FILING DATE: 26-Apr-2001  
CLASSIFICATION DATA:  
PRIOR APPLICATION DATA:  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 08/846,017  
APPLICATION NUMBER: US/08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US/08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
/note= "human telomerase reverse transcriptase (hTERT) catalytic protein component"  
SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-09-843-676-224  
Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 10 Gaps: 13  
US-09-042-460-2 (1-1122) x US-09-843-676-224 (1-4015)  
Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATCCGCGGCTCCCCGCTGCGAGCGCTGCTCTCTGCTGCGGAGCCACTACCGCAG 115  
Qy 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGGCCACGTTCGTGCGCGCTGCGGCGCCCGCCAGGCTGCGTGTGTCG 175  
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGACCCCGCGGCTTTCGCGGCTGTGTGGCCAGTGTCTGTGTGTGTCCTGG 235  
Qy 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GACGACGCGCGCGCCCGCCCGCCCTCTCTCCGCGAGGTGCTCTGCTGTAAGAGAGCTG 295  
Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100

b 296 GTGGCCCGAGTCTGCAGAGGCTGTGCAGCGCGCGGCGAGAACGTCGTGGCTTCGGC 355  
y PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTCCGGCTGCTGCAGCGGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
y SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACTGCTCCCAACACGGTGACCGACGACGTCGCGGGGAGCGGGGTGGGGCTGCTG 475  
y LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 GTGGCGCGCGTGGCGGACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 535  
y LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGGTGGCTCCAGCTGTGCGCTTACCAGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCT 595  
y ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCCACTCAGCGA 649  
y AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TGGGAGCGCGCGCTGGAACCATAGCGTCAGGAGCGCGCGG 688  
y LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
b 689 GTCCCGCTGGGCTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 748  
y ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259  
b 749 CTCCTGCTCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 808  
y -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
b 809 GCGCAGGGGTCTTGGCGCCACCGCGCGAGCGCGCTGGNCCGAGTGACCGTGTCTGT 868  
y ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
b 869 GTGGTGTCTACCTCCAGG-----CCCGCGGAGAGCCACCTCTTTGGAG 913  
y GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311  
b 914 GTGGCGCTCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
y SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCCCACATCCACATCGCGCGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTGTACGCC 1033  
y GlnThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACATCTCTACTCTCAGCGGAC---AAGGAGCAGCTGGCGCCCTTCCTTC 1090  
y LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTGTAGCGCCACGCTGACTGGCGCTCGGAGCTCGTGAGACCATCTTT 1150  
y LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
b 1151 CTGGGTTCAGCGCGCTGGATCCAGGAGCTCCCGCAGGTTGCCCGCGCTGCCCGCGCGC 1210  
y TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
b 1211 TACTTGGCAATCGCGCCCTGTCTTCTGGAGCTGTGGGACACCGCGCGAGTGCCCTTAC 1270  
y ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
b 1271 GGGGTGCTCTCAGAGCGCACTGCCGCTCGGAGCTGCG-----GTACCCCGCAGCA 1321  
y Leu-----Asn 432  
b 1322 CCGCGTGTCTGTGCGCGGAGAGCCCGCGGCTCTGTGGCGCGCGCGCGCGCGCGAGGAC 1381

QY 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
Db 1382 ACAGACCCCGCTCGCTGT 1441  
QY 453 GlyPheLeuArgAlaCysLysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTCTCGTGGGGCTCGCTGT 1501  
QY 473 AsnGluArgArgPhePheLeuLeuLysPheLeuLysPheLeuLysPheLeuLysPheLeu 492  
Db 1502 AACGAACCGCGCTTCTCAGAAACACCAAGAGATTCTTCTCTGTGGGAGAGCATGCCAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTCAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGCGCAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCCGGGGTGGCTGT 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1682 AAGTTCCTGCATCGCTGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTCCAGGAGACACGCTTTCAGAAACACAGGCTCTTTTCTACCGGAGAGGTGTGTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AGTTGTCAAGCATTTGNAATCAGACACACTTGAAGAGGGTGCAGCTGCGGGAGCTGTGC 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTGAGGAGCATCGGGAAGCAGCGCGCGCGCTGTGTGTGTGTGTGTGTGT 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCTATCCCAAGCTGTACGGCTGTGGCGCATGTGTGAACATGTGACTACGTGTGTGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAACGTTCCGACAGAAAGAGCGCGCGCGCTCTACCTCGAGGGTGAAGGACCTGTTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCAACTACGAGCGCGCGCGCGCGCGCGCGCTCTGTGGCGCTCTGTGTGTGGC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGACGATATCCACAGCGCGCTGGCGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGTACTTTGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2221  
QY 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGGCTCAGGAGGTCTATCGCCAGCATCATCAA---CCCCAGAACAGTACTGTGGTG 2278  
QY 733 ArgGlnTyrAlaValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
Db 2279 CGTCGGTATGCGGT 2338  
QY 753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2339 CAGCTCTTACCTTGCAGAGCTCCAGCGGTACATGGACAGTTCGTGGCTCACCTCAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2399 GAGACC-----AGCCCGCTGAGGATCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2452



Qy	312	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle	330
Db	974	CCCCCATCCACATCGCGGCCACACACATCGCTCCGGACACGCCCTGTGCCCCGGGTGATGCC	1033
Qy	331	GluThrArgHisPheLeuTyrsrArgGlyAspGlyGlnArgLeuAsnProSerPhe	350
Db	1034	GAGACCAAGCACTTCCTACTCTCTCAGGCGAC---AAGGAGCAGCTGGGCGCTCTTC	1090
Qy	351	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe	370
Db	1091	CTACTCAGCTCTCTGAGGCCGACCTGACTGGCGCTCGGAGCTCGTGAGACCATCTTT	1150
Qy	371	LeuGlySerArgProArgThrSerGlyProLeuCyAsrThrHisArgLeuSerArgArg	390
Db	1151	CTGGGTTCAGGCCCTCGATGCCAGGAGCTCCCGCAGGTTCCTCCCGCTGCCCCAGGCGC	1210
Qy	391	TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr	410
Db	1211	TACTGGCAATGCGGCCCTCTGTTCTGAGCTCTTGGGAACACGCGCAGTGCCCTTAC	1270
Qy	411	ValArgLeuLeuArgSerHisCyAsrPheArgThrAlaAsnGlnValThrAspAla	430
Db	1271	GGGGTGCTCTCAAGACGCACTGCCCGCTCGAGCTGCG-----GTCAACCCAGCA	1321
Qy	431	Leu-----Asn	432
Db	1322	GCGGTGCTGTGCCGGGAGAACCCACGGGCTGTGTGGCGGCCCGCAGGAGGAGAC	1381
Qy	433	ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr	452
Db	1382	ACAGACCCCGTGGCTGGTGACGTCTCCGCGACACACAGACGCCCTTGGCAGGTGTAC	1441
Qy	453	GlyPheLeuArgAlaCysLeuCyLeValValValSerAlaSerLeuTrpGlyThrArgHis	472
Db	1442	GGCTCGTGGCGGCTGCTCGCGCGCTGGTCCCCCAGGGCTCTGGGGCTCCAGGAC	1501
Qy	473	AsnGluArgArgPhePheLeuAsnLeuLysPheIleSerLeuGlyLysTrpGlyLys	492
Db	1502	NACGAACGCCGCTTCTCAGGAACACCAAGAAATTCATCTCCCTGGGAAGCATGCCAAG	1561
Qy	493	LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer	512
Db	1562	CTCTCGCTCAGAGACTACGTGAAGATCAGCGTGGGACTTGCCTTGGCTGGCGCAGG	1621
Qy	513	SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla	532
Db	1622	AGCCAGGGGTGGCTGTGTTCCGCGCGCAGACACCGTCTCGTGAGGAGATCTCTGGCC	1681
Qy	533	ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr	552
Db	1682	AGTTCCTGCATCTGCTGATGAGTGTAGCTGTGAGCTGCTCAGGTCTTCCTTTTAT	1741
Qy	553	IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer	572
Db	1742	GTCAAGGAGACCCAGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC	1801
Qy	573	LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer	592
Db	1802	AGTGTCAAAGCATTTGGAAATCAGACACATTTGAAGAGGTGCAGCTGCCGAGGTGTGCG	1861
Qy	593	GlnGluValValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg	612
Db	1862	GAAGCAGAGGTCAAGCAGCATCGGGAAGCCAGGCGCGCGCTGTGCAGCTCCAGACTCCGC	1921
Qy	613	PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr	632
Db	1922	TTATCCCCAGCCCTGACGGGCTCGGCCCATTTGTGAACATGACTACGTCTGGGGAGCC	1981
Qy	633	ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrClnArgLeuLysThrLeuPhe	652
Db	1982	AGAACCTTCCGCAGAAAGAGGGCCGAGCGTCTCACTTCAGAGGTGAAGCACTGTTC	2041



653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGGTGCTCAACTACGAGCGGGCGCGCGCCCTCTCTGGGGCTCTGTGTGGGC 2101  
673 MetAsnAspTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGGCTGGCGACCTTCTGCTGCTGCTGGCGCCCGAGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaTleProGln 712  
2162 CCGCTGAGCTGACTTGTGCAAGGTGGATGTACGGGGCGGTAGACACCATCCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTACGGGAGGTATCGCCAGCATCATCAAA---CCCAAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGGTATGCGGTGGTCCAGAAGCGCGCCATGGGCACGTCCGAAGGCCCTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACCTTGACAGACCTCCAGCGGTACATCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGCTGCTCATCGACAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLysLys 812  
2453 GAGCCAGACGATGGCTCTTCGACGCTTCTCTACGCTTCATGTGCCACACCGCGTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTACGTCCAGTGCAGGGATCCGAGGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCACCTGTGCTACGGGCAATGAGAACAGCTGTTCCGGGGATCGGCG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCGCTGCTGGTGGATGATTTCTTGTGTGACACCTCACCTACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 CGCAAAACCTTCTCAGACACCTGCTGTCGAGGTGTCCTGAGTATGGCTGCGTGTGAAC 2752  
893 LeuGlnLysThrValAlaAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGGGAGACAGATGGTGAATCTCCCTGTAGAACAGAGGCCCTGGGTGGCAGGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGGCCACGCGCTATTTCCTGCTGGCGCTGCTGTGTGATACCGGAC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGTCAGACGCTACTTCAGTATGCCGGACCTCCATCAGACCGAGTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGGAGAACATCGTCGCAAACTTTTGGGTCTTCGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACACCTGTTCTGATTTGAGTGAACAGCTCCAGACGGTGTGACCC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAGATCTCTCTGCTCAGCGGTACAGGTTTTCACGCATGTGTGCTGCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032

3113 CCATTTCATCACAAGTTTGGAGAACCCACATTTTCTGCGCGTCATCTCTGACACG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCTCTCCTCTGTCTATCTCATCTCTGAAGCCAGAACGAGGATGTGCTGGGGCCCAAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3233 GGGCGCGCGCGCTCTGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGTGTCACTACGTGCCACTCTCTGGGTCCTCAGACGACCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACCGACTGAGTCGAGAGCTCCGGGAGGAGCTGCTGCTGGAGCTGCTGGAGCGCGCAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3451

## RESULT 6

US-10-053-758-224  
; Sequence 224, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1e1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:

QY 241 ValProSerAlaIysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
DB : : : : :  
QY 749 CTGGCGTTGCCAAGAGCCACAGCTGCCTGCCCTTGACGCGAGCGACGCCCGTT 808  
DB : : : : :  
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
DB GGGCAGGGGTCTGGGCCCAACCGCGGCAGGACGCTGCAGCACGATGACCGTGGTTCTGT 868  
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
DB : : : : :  
QY 869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGAACCCACCCTCTTTGGAG 913  
DB : : : : :  
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLeuHisLysPheSer 311  
DB : : : : :  
QY 914 GGTGGCTCTCTGCACGCGCCACTCCACCATTCGTGGCGCCGACACACCGCGGC 973  
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
DB CCCCATCCACATCGCGGCCACACAGCTCCCTGGGACACGCTTGTCCCCTGGTACGCC 1033  
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
DB : : : : :  
QY 1034 GAGCAACAAGCACTCTCTACTCTCAGCGAC---AAGGACGAGCTGCGGCCCTCTTC 1090  
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
DB : : : : :  
QY 1091 CTACTCAGCTCTCTGAGGCCACGCTGCTGCGCTCGAGGCTCGTAGACCATCTTT 1150  
QY 371 LeuGlySerArgProArgThrSerdGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
DB : : : : :  
QY 1151 CTGGGTTCCAGGCCCTCGARGCCAGGAGCTCCCCGAGGTGCCCGCTCCCGCAGCGC 1210  
QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
DB TACTGGCAAATGCGGCCCTCTTTCTGGAGCTGCTTGGGAACACACGCGAGTGCCTTAC 1270  
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
DB GGCGTGTCTCTCAAGAGCACTGCGCGCTGCGAGTGC-----GTCAACCCAGCA 1321  
QY 431 Leu-----Asn 432  
DB GCGGTGTCTGTGCCGGGAGAGCCCGAGGCTCTGTGGCGGCCCGGAGGAGGAC 1381  
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
DB ACAGACCCCGCTGCCTGGTGCACTGCTCGGCAGACACAGCAGCCCTTGGCAGGTGATC 1441  
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
DB : : : : :  
QY 1442 GCCTTGTGCGGGCTGCTGCGCGGCTGGTCCCGCCAGGCTCTGGGCGCTCCAGGAC 1501  
QY 473 AsnGluArgArgPhePheLeuAsnLeuLysPheIleSerLeuGlyLysTyrGlyLys 492  
DB ACGAACCGCGCTTCTCAGGAACACCAAGAATTCTATCTCCCTGGGAAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
DB CTCCTGCTCAGGAGCTGACCTGGAAGATCAGCGTGGGACCTGGCGCTGGCGCGCAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
DB AGCCCAAGGGTGGCTGTGTTCCCGCCGCGACAGACACCGTCTGCGTGAGAGATCTGSCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
DB AAGTCTCTGCACTGGCTGATGAGTGTAGCTCGTCGAGCTGCTCAGGTCTCTTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
DB : : : : :  
QY 1742 GTCAGGAGACACAGCTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
OY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592

```
1802 AAGTTGCAAGCATTTGATCAGACGACCTTGAAGAGGGTGCAGCTGGGGAGCTGTGC 1861
593 GlnGluGluValAlaArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612
1862 GAACAGAGGTTCAGCAGCATCGGAGACCCAGGCGCCCTGCTGACGTCTCAGACTCCGC 1921
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTrpSerMetGlyThr 632
1922 TTATCCCAAGCTGACGGGCTGCGGCGGATTTGGAATGAACTAGCTGCTGGGAGCC 1981
633 ArgAlaLeuGlyArgArgGlyGlnAlaHisPheThrGlnArgLeuLysThrLeuPhe 652
1982 AGAAGCTTCCGAGAGAAAGAGGCGGAGCGCTCACCTCGAGGGTGAAGGCACTGTTTC 2041
653 SerMetLeuAsnTrpGluArgThrLysHisPheProHisLeuMetGlySerValLeuGly 672
2042 AGCGTGCTCAACTACGAGCGGCGCGGCGCCCTCTCTGGCGGCTCTGTGCTGGGC 2101
673 MetAsnAspIleTrpArgThrTrpArgAlaPheValLeuArgValAlaLeuAspGln 692
2102 CTGGACGATATCCACAGGCGCTGCGGCACTTCTGCTGCTGCTGGGCGCCAGGACCG 2161
693 ThrProArgMetTrpPheValLysAlaAspValThrGlyAlaTrpAspAlaIleProGln 712
2162 CGCGCTGAGCTGTACTTGTCAAGGTGGATGTGACGGGCGCTACGACACCACTCCCGCAG 2221
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTrpCysIle 732
2222 GACAGGCTCAGGAGGTTCATGCCAGCATCATCAA---CCCGAGACAGCTACTCGTG 2278
733 ArgGlnTrpAlaValAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2279 CGTCGTATCGCTGTGTCAGAGCGGCCCATCGGCACCTCGCGCAAGCTTCAAGAGC 2338
753 GlnValThrThrLeuSerAspLeuGlnProTrpMetGlyGlnPheLeuLysHisLeuGln 772
2339 CAGTCTCTACTTGACAGACCTCAGCGGTACATGCGACAGTTCGTGGCTACCTCGAG 2398
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2399 GAGACC-----AGCCCGCTCAGGAGTCCGCTCATCGACGACAGCTCCTCCCTGAAT 2452
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812
2453 GAGCGCAGCAGTGCCTCTCGAGCTCTCAGCTTCATCGCTTCATGTCACACACCGCGCGC 2512
813 IleGlyAspArgCysTrpThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2513 ATCAGGGCAAGTCTAGTCCAGTGCAGGGGATCCGCGAGGGCTCCATCCTCTCCAGC 2572
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2573 CTGCTCTCAGACCTGTGTACGGGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGG 2632
853 AspGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 872
2633 GACGGGCTCTCTCGCTTGGTGGATGATTTCTGTTGGTGACACTCACCTCACCCAC 2692
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTrpGlyCysMetIleAsn 892
2693 GCGAAACCTTCTCCTCAGGACCTGCTCGAGGTGTCTCCTGATGCTGCTGGTGAAC 2752
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912
2753 TTGCGGAGACAGTGGTGAATCTCCCTGTAGAGACGAGGCGCTGGGTGGCAGCGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATGCGGCCCGCCAGGCTTATTCCTCTGGTGGGCTGTGCTGTGATACCCGAC 2872
933 LeuGluValPheCysAspTrpSerGlyTyAlaGlnThrSerIleLysThrSerLeuThr 952
```

```
Db 2873 CTGAGGTGCAGAGCGACTACTCCAGCTATGCCGACCTCCATCAGAGCCAGTCTCACC 2932
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972
Db 2933 TTCACCGCGGCTTCAAGGCTGGAGAAACATGCGTCGCAAACTCTTTGGGCTCTTGGG 2992
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
Db 2993 CTGAGGTGCACAGCTGTTTCTGGATTTCAGGTGAACAGAGCTCCAGAGCGGTGACACC 3052
Qy 993 AsnIleTrpLysIlePheLeuGlnAlaTrpArgPheHisAlaCysValIleGlnLeu 1012
Db 3053 AACATCTACAGATCTCTCTGCTGCGAGGCTACAGGTTTTCACGCACTGTGTGTCAGCTC 3112
Qy 1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3113 CCATTCATCAGCAAGTTTGAAGAACCCCACTTTTCTCGGGCTCATCTCTGACAGC 3172
Qy 1033 AlaSerCysCysTrpAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
Db 3173 GCCTCCCTCTCTACTCCATCTCTGAAAGCAAGACGAGGATGCTGCGGGGCGCAAG 3232
Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTrpGlnAlaPheLeu 1069
Db 3233 GCGCGCGCGCGGCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3292
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTrpLysCysLeuLeuGlyProLeuArgThrAla 1089
Db 3293 CTCAGGTGACTCGACACCGTGTCACTACGTGCCACTCTCTGGGTCACCTCAGGACGCC 3352
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
Db 3353 CAGACGAGCTGAGTCGGAAGCTCCCGGGGAGCAGCGTCACTGCGCTGAGGCGCGAGCC 3412
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3413 AACCGGCACTGCCCTCAGACTTCAGACATCTCTGGAC 3451
```

## RESULT 7

```
US-10-208-243-1
; Sequence 1, Application US/10208243
; Publication No. US20030044394A1
; GENERAL INFORMATION:
; APPLICANT: Gaeta, Federico C.A.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune
; TITLE OF INVENTION: Response to a Telomerase Antigen
; FILE REFERENCE: 015389-003500PC
; CURRENT APPLICATION NUMBER: US/10/208,243
; CURRENT FILING DATE: 2002-07-30
; PRIOR APPLICATION NUMBER: US/09/675,321
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/112,006
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(3454)
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)
US-10-208-243-1
```

```
Alignment Scores:
Pred. No.: 0 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
```

Query Match: 59.40% Indels: 52  
B: 15 Gaps: 13  
S-09-042-460-2 (1-1122) x US-10-208-243-1 (1-4015)  
Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
b 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAG 115  
Y 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
b 116 GTGCGCGCGCTGCGCAGCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCTGCGCGAG 175  
Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
b 176 CCGCGGAGCCCGCGCGCTTCCCGCGCTGCTGCGCGCGCTGCGCGCGCTGCGCGCTG 235  
Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerLeuLysGluLeu 80  
b 236 GAGCGAGCTG 295  
Y 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
b 296 GTGCGCGAGTCTGCGAGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGC 355  
Y 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTCGCGCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGC 415  
Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACCTCCCAACACGCGTGACCGACACACTGCGCGCGCGCGCGCGCGCGCGCTG 475  
Y 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 CTGCGCGCGCTGCGCGAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 535  
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGCGTGGCTCCAGCTGCGCTACCAGGTGTGCGCGCGCGCGCTGTACAGCTCGCGCT 595  
Y 181 ThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCACCTACG 649  
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TCCGACGCGCGCTGGAACCATAGCGTCAGGAGCGCGCG 698  
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerSer 240  
b 689 GTCCCTCTGGCTGCGACCGCGCGTGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748  
Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTGCGCTGCGCCAGAGCGCGCGCGCGCTGCGCTGCGCTGCGCGCGCGCGCGCGCGTT 808  
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
b 809 GGCAGCGGCTCTGGCGTGT 868  
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
b 869 GTGCTGTACCTGCCAGA-----CCCGCGGAGAGCGCACCTCTTTGGAG 913  
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysHisLysPheProSer 311  
b 914 GGTGCGCTCTGTGCGACGGCGCGCTCCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCG 973  
Y 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---Phele 330  
b 974 CCCCGATCCACATCG 1033

331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACCTTCTCTACTCTCTCAGCGAC---AAGAGACAGCTGGCGCTCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCAGCTGACTGGCGCTCGGAGGCTGCTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTTCCAGGCGCTGGAGCCAGGAGACTCCCGCAGGTGCTCCCGCTGCGCCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGGCGCCCTGTTTCTGGAGCTGCTTGGGAACCAAGCGCAGTGGCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGTGTCTCTAAGAGCGCACTCCCGCTGCGAGCTGCG-----GTACCCCGAGCA 1321  
431 Leu-----Asn 432  
1322 CCCGCTGTCTGTCCCGGAGAACCCCGAGGCTCTGTGGCGCGCCCGCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCGCTGTGTCAGCTGCTCCGCGCAGCAGCAGCGCCCTGGCAGGTGAT 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GCTTGTGCGCGCTGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTG 1501  
473 AsnGluArgArgPhePheLeuLeuLysLysPheLeuSerLeuGlyLysTyrGlyLys 492  
1502 AACGACCGCGCTCTCTCAGGAAACACAGAGATTCATCTCTCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCGCTGCGAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTGGCTGGCGCAG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGTGTGCTGTCTCGCGCGCAGAGCACCGTCTGCTGAGGAGATCTCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTCTCTGCACTCGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741  
553 IleThrGluSerThrPheGlnLysAsnArgPhePheTyrArgLysSerValTrpSer 572  
1742 GTCAGGAGACACGTTTCAAAAGACAGAGCTCTTTTCTACCGGAAGATGTCTCGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTGAATCAGACACTTAAGAGGGTGCAGCTGGGAGGAGTGTGCG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGTTCAGGAGCATCGGAAGCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTGAGCGGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981  
633 ArgAlaLeuGlyArgArgGlnAlaGlnHisPheThrGlnArgLeuLeuValThrLeuPhe 652  
1982 AGAAGTTCCCGACAGAAAGAGGCGCGAGCTCTACCTCGAGGGTGAAGGACACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCTCAACTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuArgGln 692

2102 CTGGACGATATCCACAGGCTGGCGGACCTTCGTGCTGTGGCGGCCAGGACCCG 2161  
693 ThrProArgMetTyrPheValIleAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTACTTGTCAAGTGATGTGACGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTATCGCCAGCATCATCAA---CCCGACAGACAGTACTCGGTG 2278  
733 ArgGlnTyrAlaValAlaArgAspSerGlnGlyGlnValHisLysSerPheArgAsp 752  
2279 CGTCGGTATCGGTGTCCAGAGGCCGCCATCGGACGTCGCGAGGCGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGGCTCTACCTTGACAGACTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCCCTGAGGATGCCGTCTCATCGACAGAGCTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGCAGCGTGGCTCTTCGAGCTCTTCATCGCTTCATGTCGCCACACCGCGGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCGAAGTCTCTACGTCAGTGCAGGGGATCCCGACAGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGCCTGTGTACGGCGACATGAGAACAAAGCTGTTTGGCGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCCTCGTGTGGTGATGATTCTGTGTGGTGCACACCTCCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GGGAAACCTTCTCAGACACCTGTGCGAGGTGTCTCTGATGATGCTGGTGGTGAGC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGAAGACAGTGGTGAATTCCTCTAGAACAGCAGGCGCTGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCCCGGCCACCGCTATTCCTCGGTGGCGCTGTGCTGGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTTGACAGGCGACTCTCCAGCTATGCGCGGACCTCCATCAGAGCGAGTCTCAC 2932  
953 PheGlnSerValPheIleAlaGlyThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGGCTTCAGGCTGGGAGAACATGGCTCGCAACTCTTTGGGGCTTTCGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAGTGTACAGCTGTTCTGGATTGAGGTGAGGTGAACAGCTCCAGAGCGGTGTGCACC 3052  
993 AnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAGATCTCTCTGTGTCAGCGGTACAGGTTTTCACGATGTGTGCTGCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCACAGAGTTGGAGAACCCACATTTTTCGCGCGTCACTCTGCACAGC 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052

3173 GCCTCCCTGTGCTACTCCATCTGTAAGAACGACAGGAGGATGTCTGCTGGGGCCAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GGCGCCGCGGCTCTGCTCCCTCGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTGACACACCTGTACCTAGTGCCTCTGCTGCTGCTGCTGCTGCTGCTG 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3353 CAGACGCGAGTGTGAGTGGAGCTCCCGGGGACGACCTGCTGCTGCTGCTGCTGCTGCTG 3412  
1110 AspProLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCACTGCCTCAGACTTCAAGACCATCTCGAC 3451

RESULT 8  
US-10-054-295-224  
Sequence 224, Application US/10054295  
Publication No. US20030044953A1  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030044953A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,295  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0300  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"  
 /notes "human telomerase reverse  
 transcriptase (hTRT) catalytic protein  
 component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

0-054-295-224

ment Scores:

1. No.: 0 Length: 4015  
 e: 3505.00 Matches: 719  
 ent Similarity: 72.94% Conservative: 122  
 Local Similarity: 62.36% Mismatches: 260  
 y Match: 59.40% Indels: 52  
 Gaps: 13

19-042-460-2 (1-1122) x US-10-054-295-224 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCCCGCGCTCCCGCTGCCGAGCGTGGCTCCCTGCTGGCAGCCACTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
 116 GTGCTCCGCTGCCACGTTCTGCGGCGCTGGGCGCCCGAGGCTGGCGCTGGTCAG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGGGGGAGCCCGCGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGGCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GAGCGACGGCGCGCCCGCCCGCTCTTCGCGCAGGTGCTGCTGCAAGAGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGGCCGAGTGTGCAGAGGCTGTGCGAGCGCGCGCGAGAGAGCTGCTGGCTTCGGC 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCTGTGTGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGTACCTGCTCCCAACACGCTGACCGACCTGCGGGGAGCGGGGCGTGGGGCTGCTG 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 476 CTGGCGCGCTGGGCGACGACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 536 CTGTGCTGCCAGCTGCCCTACCAAGTGTGGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 595  
 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 596 GCGACTCAGCG 649  
 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 650 -----TGCAGACGGCGCTGGACCATAGCGTCCAGGAGCGCGCG 688  
 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 689 GTCCCCCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAAGT 748  
 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 749 CTGCGGTGCCAGAGCTT 808  
 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 809 GGGCAGGGGTCTGGGCTTCTGT 868

QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 DB 869 GTGTGTGTCCTGCCAGA-----CCGCGCGAAGAACGCCACCTCTTTGGAG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 914 GGTGGCTCTCTGGCAGCGCGCACTCCACCCATCGTGGCGCGCCAGCACCGCGGGC 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 974 CCCCACATCCACATCGCGGCCACCACTGCTGGGACACCGCTTGTCCCCGGGTGTACGC 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 1034 GAGACCAAGCACTTCTCTACTCTCAGGCGAC---AAGGAGCAGCTGGCGCTCTCTTC 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 DB 1091 CTACTCAGCTCTCTGAGCGCCAGCTGACTGGCGCTCGGAGGCTCGTGAGACCATCTTT 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 DB 1151 CTGGGTTCAGGCGCTGATGCCAGGACTCCCGCAGGTTCGCCGCTGCCCCAGCGC 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1211 TACTGGCAATCGCGCCCTGTTTCTGGAGTGTCTGGGAACCAACGCGCAGTGCCTTAC 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 DB 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTCGAGCTGCG-----GTACCCCGACGA 1321  
 QY 431 Leu-----Asn 432  
 DB 1322 GCCGTGTCTGCGCGCGAGAACCCCGAGGCTCTGTGGCGCCCGCGAGGAGGAC 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 DB 1382 ACAGACCCCGCTGCGCTGTGTGACGCTCTCGCGCAGCACAGCAGCCCTGGCAGGTGTAC 1441  
 QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 DB 1442 GGTCTCGTGGGCGCTGCTGCGCGGCTGTGTGCGCCCGCCCGCTCTGGGCTCCAGGCAC 1501  
 QY 473 AsnGluArgArgPhePheLeuAsnLeuLysPheLeuSerLeuGlyLysTyrGlyLys 492  
 DB 1502 AACGAACCGCTTCTCTCAGGAACACCAAGAAAGTTCTCTCCCTGGGGAAGCATGCCAAG 1561  
 QY 493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 DB 1562 CTCTCGTGCAGAGCTGACGTGGAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAG 1621  
 QY 513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgGileLeuAla 532  
 DB 1622 AGCCAGGCGTGGCTGTGTTCCGCGCGCAGACGCGCTCTGCGTGAGAGATCCTGGCC 1681  
 QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 DB 1682 AAGTCTCTGCACTGGCTGATGAGTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1741  
 QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
 DB 1742 GTACCGGAGCACCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC 1801  
 QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
 DB 1802 AAGTTGCAAGCAATGGAAATCAGACGACTTGAAGAGGGTGCAGCTGCGGAGCTGTGC 1861  
 QY 593 GlnGluValValArgHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
 DB 1862 GAAGCAGAGTTCAGCAGCATCGGAAGCCAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 1921



613 PheileProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCCCTGACGGGCTGCGCGCATGTGAACATGGACTACGTCGTGGAGCC 1981  
633 ArgAlaLeuGlyArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGGCGCGGAGGCTCTCACTCGAGGGTGAAGGACCTGTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCAACTAGCAGGCGCGCGCGCGCGCTCTCTGGCGCTCTGTGCTGGCG 2101  
673 MetAsnAspIleTyrArgThrPheArgAlaPheValLeuArgValAlaLeuAspGln 692  
2102 CTGGAGCATATCCAGAGGCTGGCGGCACTTGTGCTGCTGCTGGCGCGCGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGACTTGTCAAGGTGATGTACGGCGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGCTCACGAGGTATCCCGAGCATCATCAAA-----CCCCAGACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTGGGTATGCCGTGCTCAGAAAGCGCCCATGGGCACGTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTTACCTTGACAGACCTCCAGCGGTACATCGACAGTTCGTGGGTCTCCTCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCGCTGAGGATGCGGTGTCATCGAGAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGCAGGTGGCTCTTCGACGCTCTCTTACGCTTCATGTGCCACCAAGCGCGTCCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTTACCTCAGTCCAGGCGGATCCCGGAGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACGCTGTGCTACGCGGCATGAGAAAGAGTGTTCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTGCTGTTGGTGGATGATTTCTTTGGTGGACACCTCACCCTCACCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTCGTCCAGGTGTCCTGAGTAGTGGTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGGTGAATCTTCCCTGTGAGAAAGCAGGCGCTGGGGGACCGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrPysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGCCACCGCTATTTCCTGCTGCGGCTGCTGCTGATACCCGAC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGCAGGCACTACTCCAGTATGCCGACCTCCATCAGAGCCAGCTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGGCTTCAAGGCTGGAGGAACATGCTCGCAACTCTTTGGGGTCTTGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992

Db 2993 CTGAAGTGTACAGCCCTGTTTCTGAAATTCAGGTGACAGCTCCAGACGTTGTGCACC 3052  
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGCTGCGAGCGGTACAGGTTTTCACGATGTGTGTGCAGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTCATCAGCAAGTTTGGAAAGACCCACATTTTCTCGCGCTCATCTCTGCACAG 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCCTCTGCTACTCCATCTGAAAGCCAGACGAGGATGTGCTGGGGGCGCAAG 3232  
QY 1053 GlySer-----PheProProGluAlaAlaHisTyrPleuCysTyrGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCCCTCTGCGCTCCGAGGCGGTGACGTGGCTGTGCCACCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAGCTGATCTCCAGACCGGTGTACCTAGTCCACTCTCTGGGGTCACTCAGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
Db 3353 CAGACGCGAGCTGAGTCCGAAAGCTCCCGGGGACGACGTGACTGCTGGAGGCCGAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGACATGCGCTCAGACTTCAGACCATTCCTGGAC 3451

## RESULT 9

US-10-054-611-224  
Sequence 224, Application US/10054611  
Publication No. US20030059787A1  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030059787A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,611  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

## INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"

/note= "human telomerase reverse

transcriptase (hTRT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

-10-054-611-224

## Alignment Scores:

Seq. No.:	0	Length:	4015
Score:	3505.00	Matches:	719
Percent Similarity:	72.94%	Conservative:	122
Local Similarity:	62.36%	Mismatches:	260
Identity Match:	59.40%	Indels:	52
Gap:	15	Gaps:	13

-10-054-460-2 (1-1122) x US-10-054-611-224 (1-4015)

```
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGGCGGCGCTCCCTCCCTCCAGCGCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluArgLeuValGln 40
116 GTGCTGCGCTGGCCAGCTTCTGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethIle 60
176 CGCGGGACCGCGCGCTTTCGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCT 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
236 GACGACGCGCGCGCGCGCGCGCGCGCTTCCGCGCGCTGGCGCGCGCTGGCGCGCT 295
81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyProProMetAlaPheThrSerValArg 120
356 TTCGCGCTGTGGACGGGCGCGCGCGCGCGCGCGCGCGCTTCCACCGCGCGCTGGCG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
416 AGCTACTCTGCCCAACCGGTGACCGCGCTGGCGCGCGCGCGCGCGCTGGCGCGCT 475
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCGCGCTGGCGCG 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGTGTGCTCCAGCTGCGCTACCGGTGCGCGCGCGCGCGCGCGCTGTACCGCTCG 595
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCCACTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
```

553 IleThrGluSerThrPheGlnIysAsnArgLeuPhePheThrArgLysSerValTrpSer 572  
1742 GTTACGGAGACCAAGCTTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGlnArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCAATTGGAAATCAGACAGCACTTTGAAGAGGTGCGAGCTGCGGAGCTGCG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTGAGCAGCATCGGAGAACCGAGCGCGCTGTGAGCTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTrpSerMetGlyThr 632  
1922 TTTCATCCCAAGCTTGACGGCTGCGGCGGATTTGAACTGACTAGCTCGTGGAGGCC 1981  
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCCGAGAGAAAGAGGCGCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTrpGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 ACGGTGCTCACTAGCAGCGGCGCGGCGCGCGCTCTGTGGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTyArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGCGCTGGCGACCTTGTGTGTGGTGGCGGCCAGGCCG 2161  
693 ThrProArgMetTyPheValValAlaAspValThrGlyAlaTyArgAlaIleProGln 712  
2162 CGGCTGAGCTGTACTTTGTCAAGTGTGATGACGGCGCGGTACGACACCATCCCGCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGlnSerThrTrpCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGACACGTACTGGTG 2278  
733 ArgGlnTyAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGTATGCGTGTGTCAGAAAGCGCGCCATGGGCACGCTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyTrpMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACCTTGACAGACTCCAGCGGTACATGCGACAGTTCTGGTGTCTACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATCGCGTCTCATCGACGAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGCAGAGTGGCTCTTCAGCTCTTCTACGCTTCATGTCGTCACCGACCGCGTGGCG 2512  
813 IleGlyAspArgCysTyThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTCTACGTCCAGTCCAGGAGTCCGCGAGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGCTGTGTGCGGACATGAGAACCAAGCTGTTTGGGGGATTCGCGCG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTCTCTCGGTTTGGTGGATATTTCTTTGTTGTGACACCTCACCCTACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyArgCysMetIleAsn 892  
2693 GCGAAACCTCTCTCAGGACCTGTGTCCGAGGTGTCTCCAGTATGCTGCTGCTGTAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGGGAGACAGTGTGTGAATCTTCTCTGTAGAACGAGGCGCTTGGTGGCAGCGCTTTT 2812

QY 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2813 GTTCAGATGCGGCGCCACCGCTATTCCCTGTGTGGCTGTCTGTGTATACCCGAGCC 2872  
QY 933 LeuGluValPheCysAspTyTrpSerGlyTyArgAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2873 CTGGAGGTGCGAGAGCACTACTCCAGCTATGCCCGACCTCCATCAGACCGAGTCTCACC 2932  
QY 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2933 TTCAACCGCGCTTCAAGGCTGGGAGAACATCGTCGCAAACTCTTTGGGGTCTTGGCG 2992  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2993 CTGAAGTGTACAGCTGTCTTCTGGATTTGAGTTGAGGTGAACAGCTCCAGACGCTGTGACC 3052  
QY 993 AsnIleTyLysIlePheLeuLeuGlnAlaTyArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGCTGCGAGGCGTACAGGTTTTCAGGCATGTGTGCTGACGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTTCATCAGCAAGTTTGGAAAGAACCCACACATTTTCTGGCGGTCTCTGTACAGC 3172  
QY 1033 AlaserCysCysTyArgAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaser 1052  
Db 3173 GCCTCCCTCTGTACTCCATCTCTGAAAGCCAGAACGAGGATGCTGGGGGCCAAG 3232  
QY 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyArgLysAlaPheLeu 1069  
Db 3233 GGCGCCCGCGGCTCTGCGCTCCGAGGCGGTGAGTGTGCGCCACCAAGCATTTCTGTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGCTGACTGCGACACCGTGTCTACCTGCTGCGCTCTCTGAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGAGCTGAGTCGAAAGCTCCCGGAGACGACGCTGACCTGCGCTGGAGCGCGAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGCACTGCGCTTCAGACTTCAAGACCATCTCTGAC 3451

## RESULT 10

US-10-105-963-1  
; Sequence 1, Application US/10105963  
; Publication No. US20030068818A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Denning, Chris  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
; TITLE OF INVENTION: Recombination  
; FILE REFERENCE: 731/002  
; CURRENT APPLICATION NUMBER: US/10/105,963  
; CURRENT FILING DATE: 2002-03-21  
; PRIOR APPLICATION NUMBER: US 60/277,811  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION:  
US-10-105-963-1

## Alignment Scores:

```

ed. No.:      0      Length:      4015
ore:          3505.00  Matches:      719
rc:           72.94%  Conservative: 122
st Local Similarity: 62.36%  Mismatches: 260
ery Match:     59.40%  Indels:      52
:              15      Gaps:       13

-09-042-460-2 (1-1122) x US-10-105-963-1 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATCCGCGCGCTCCCGCTCCGCGAGCGGCGCTCCCTGCTGGCGAGCCACTACCGCGAG 115
21 ValTrrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGCTCCGCTGGCCACGTTGCTGGCGGCCCTGGGGCCCCAGGGCTGGCGCTGGTGCAG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrr 60
176 CGCGGGGACCGCGCGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyGluLeu 80
236 GACGCGCGCGCGCGCGCGCGCTTTCGCGCGCGCTTTCGCGCGCGCTTTCGCGCGCG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGGCCGAGTGTCTCAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrrMetLeuLeu 140
416 AGCTACCTGCTCCCAACAGGTGACCGACGACGACGCGCGCGCGCGCGCGCGCGCGCG 475
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGGCGCGCTGGCGACGAGTGTGCTGCTACCTGCTGGCAGCGTGTGCTGCTGCTGCT 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGGTGGCTCCCAAGTGTGCTACCAAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCG 595
181 ThrThrAspIleTrrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCNACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
650 -----TGCAGACGGGCTGGAACCATAGCGTCAAGGAGCGCGCGCGCGCGCGCG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240
689 GTCCCTCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
749 CTGCGGTGTCAGAGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrrPhe 272
809 GGGCAGGGTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
869 GTGGTGTACCTGCCAGA-----CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 913
293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311
914 GGTGCGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973

```

```

312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro----PheIle 330
974 CCCCACATCCACATCGCGGCCACACACCGCTTGTGCGGACACGCTTGTCCCCGGGTACGCC 1033
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
1034 GAGACCAAGCACTTCTCTACTCTCTCAGGGGAC---AAGAGGAGCTGGCGCTCTCTTC 1090
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
1091 CTACTCAGCTCTCTGAGGCCCGACCTGCTGGAGCTGCTTGGGAACACACGCGCAGTGCCTTT 1150
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
1151 CTGGGTTTCCAGGCTTGGATGCCAGGACTCCCCGAGGTGGCCGCTTGGCCGCGCGC 1210
391 TyrTrrPginMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
1211 TACTGGCAATGCGGCCCTTGTTCGGAGCTGCTTGGGAACACACGCGCAGTGCCTTAC 1270
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGCGAGCTGCG-----GTCACCCACGCA 1321
431 Leu-----Asn 432
1322 GCCGCTGTCTGCGCGGAGAACCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAC 1381
433 ThrSerProProHisLeuMetAspLeuLeuArgHisSerSerProTrrPginValTyr 452
1382 ACAGACCCCGCTGCGCTGTCAGCTCTCGCGCAGCACACGAGCCCTGGCAGGTGATC 1441
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrrPginTyrArgHis 472
1442 GCTTCTGGTGGGCTCTGCTGCGCGGCTGGTGGCCCGCGGCTCTGGGCTCCAGGSCAC 1501
473 AsnGluArgArgPhePheAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
1502 AACGAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAG 1561
493 LeuSerLeuGlnGluLeuMetTrrLysMetLysValGluAspCysHisTrrLysArgSer 512
1562 CTCTCGCTGCGAGAGCTGACGTGAAGATGAGCTGGGAGCTGCGCTTGGCTGGCGAGG 1621
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLleLeuAla 532
1622 AGCCGAGGGTGGCTGTGTTCCGCGCGCAGACACCGTCTGCGTGAGGAGATCCTGGCC 1681
533 ThrPheLeuPheTrrLysMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552
1682 AAGTTCCTGCACTGGCTGATGAGTGTGTGCTGCTGAGCTGCTCAGGTCTTCTTTTAT 1741
553 IleThrGluSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrrPser 572
1742 GTCCGAGAGACCGCTTTCAGAAAGACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1802 AAGTTCGCAAGCATTTGGAATCAGACAGCACTTGAAGGGGTGCAAGTGGCGGAGCTGCG 1861
593 GlnGluGluValArgHisHisGlnAspThrTrrLysAlaMetProLleCysArgLeuArg 612
1862 GAAGCAGAGGTTCAGCAGCATCGGGAACCGAGCCCGCTGCTGCTGCTGCTGCTGCTGCT 1921
613 PheIleProLysProAsnGlyLeuArgProLleValAsnMetSerTyrSerMetGlyThr 632
1922 TTCAATCCCAAGCCCTGAGCGGCTGGCGCGCATTTGTGAACATGGAATGGAATGGAATG 1981
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1982 AGACGTTCCGCGAGAAAGAGGCGCGCGCTCTCACCTCGAGGGGTGAAGGCACTGTTTC 2041

```



TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTERT"

/note= "human telomerase reverse

transcriptase (hTERT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

-10-044-692-1

Alignment Scores:

ed. No.:	0	Length:	4015
ore:	3505.00	Matches:	719
cent Similarity:	72.94%	Conservative:	122
st Local Similarity:	62.36%	Mismatches:	260
ery Match:	59.40%	Indels:	52
	15	Gaps:	13

-09-042-460-2 (1-1122) X US-10-044-692-1 (1-4015)

```

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGCGCGCGCTCCCGCGTCCGAGCGTGCCTCCCTGCTCGCGAGCCACTACCGCGAG 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGTCGCGTGCACAGTTCGTGCGCGCTGCGCGCCCGCGCGTGGCGGTGGTGGCG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
176 CGCGGGACCGCGCGCTTCCCGCGCGCTGGTGGCCAGTGCCTGCTGGTGGCGCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGlu 80
236 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGCGCGCGAGTGTGCAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGTGTGGACGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
416 AGCTACCTGCGCCACACAGGTGACCGACGCTGCGGGGCGCGGGGCGTGGGGCTGCTG 475
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGCGCGCGTGGCGGACGACGTGTGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGTGGCTCCACAGTCCAGCTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCG 595
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
650 -----TGGAAACGGCGCTGGAAACCATAGCTCAGGAGGCGCGG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240

```

```

Db 689 GTCCCCCTGGCGCTGCCAGCCCGGGTGCAGAGAGCGCGGGGCGAGTCCACGCCAAGT 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
Db 749 CTGCGGTGGCCAAAGAGCGCCAGCGGTGGCGCTGCCCTGAGCGGAGCGAGCCCGTT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
Db 809 GGGCAGGGGTCTCTGGGCCACCCCGGCGAGGACCGGTGGACCGGTGACCGTGTCTTCTGT 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
Db 869 GTGGGTGTCACTGTCAGCA-----CCGCGCGAAGAACCCACCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 914 GGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 974 CCCCCTCATCATCGCGGCCACCGCTCCCTGGGACACGCTTGTCCCGGTGTAGGCC 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACTTCTCTACTCTCTCAGGGGAC---AAGGAGCAGCTGCGCGCTCTCTTC 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
Db 1091 CTACTCAGCTCTCTGAGCGCCAGCGCTGACTGGCGCTCGGAGGCTCTGTGGAGACATCTTT 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1151 CTGGGTTCAGCGCCCTGATGCCAGGGACTCCCGCAGGTGTCCCGCGCTGCCCGCAGCGC 1210
QY 391 TyrTrpGlnMetArgProPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
Db 1211 TACTGGCAATGCGCGCTGTTCTGGAGCTGCTGGGAACACCGCGCAGTGGCCCTAC 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
Db 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGCGAGCTGCG-----GTACCCCGCAGCA 1321
QY 431 Leu-----Asn 432
Db 1322 GCCGTGTGTGTCGCCGCGAGAGCCCGAGGCTGTGGCGCGCCCGCGAGGAGGAG 1381
QY 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
Db 1382 ACAGACCCCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1441
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
Db 1442 GGTTCGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1501
QY 473 AsnGluArgArgPhePheAsnLeuLysLysPheLeuSerLeuGlyLysTyrGlyLys 492
Db 1502 AACGACCGCGTTCCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAG 1561
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
Db 1562 CTCTGCTGCGAGGCTGACGTGGAAGATGAGCGTGGCGGCTGCGCTTGGCTGCGCAGG 1621
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
Db 1622 AGCCAGGGGTGGTGTGTTCCGCGCGCAGAGCACCGTCTGCGTGAGGAGATCTTGCC 1681
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuLeuArgSerPhePheTyr 552
Db 1682 AAGTTCCTCACTGGCTGATGATGTGTACGTGCTGCGAGTGTCTCAGGTCTTCTTTAT 1741
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572

```



1742 GTCACGGAGACACAGCTTTCAAAAGNACAGGCTCTTTTCTACCGAGAGCTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAAAGCATTTGGAATCAGACAGACATTTGAAGAGGGTGCAGCTGCGGAGCTGTGC 1861  
593 GlnGluGluValArgHisIleGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACAGAGGTGAGCAGCATCTGGGAGCCAGCCGCGCTCTGTGAGCTCCACACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCCCTGAGCGGTGCGGCCGATGTGAACTAGGACTACGTGCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlyValGlnAlaGlnHisPheThrGlnArgLeuLeuPhe 652  
1982 AGAAGCTTCCGAGAGAAAGAGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrIleHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCACTACAGCGGGCGCGGCCCTCTCTGCGGCGCTCTGTGTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGAGATATCCACAGGCGCTGCGGCACCTCTGCTGTGGTGTGGCGGCCCGAGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGTACTTGTCTAAGTGTGTGACGGCGGCTGACGACCATCTCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisGlnSerThrTyrCysIle 732  
2222 GACAGGCTCAGCGAGTCTATCGCCAGCATCATCAAA---CCCGAGAACAGCTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGCTGCTCAGAGGCCCGCCCATGGCAGCTCCCGAGGCTTCAAGAGC 2338  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGTCACTCCGCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATCCGCTGCTCATCGAGCAGAGCTCTCCCTGGAAT 2452  
793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGCTGGCTCTTTCAGCGCTCTTCTACGCTTCATGTGCCACCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGCTCTAGCTCCAGTCCAGGGGATCCCGCAGGCGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGGCTGTGTACGGCGACATGAGACACAGCTGTGTGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTCTCTCGTGTGTGTGATGATTTCTGTGTGTGACACCTCACCTCACCCAC 2692  
873 AlaLeuThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTGTGTGAGGTGCTGCTGAGTATGCTGCTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAGACAGTGTGTGAACCTTCTCTGTAGAGCAGGAGCGCTGGGTGGCAGCGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACCGCTATTTCCTCTGCTGGTGGCGCTGCTGCTGGATACCGGACC 2872

QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
DB 2873 CTGAGGTGAGAGCGGACTACTCCAGTATGCCGAGCCTCCATCAGAGCCGATCTCACC 2932  
QY 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
DB 2933 TTCAACCGCGGCTTCAAGGCTGGAGGACATGCGTGGCAAACTCTTTGGGCTTTGGCG 2992  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
DB 2993 CTGAAGTGTACAGCCTGTCTTCGATTTCAGGTGACAGCTCCAGACGCTCCAGCGGTGCACC 3052  
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
DB 3053 AACATCTACAGATCTCTCTGCTGCGAGCGTACAGTTTCACGCATGTGTGCTGCAGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
DB 3113 CCATTTTCATCAGCAAGTTTGGAAAGAACCCACATTTTCTGCGCGCTCATCTCTGACAGC 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
DB 3173 GCCTCCCTCTGCTACTCACTCCATCTGAAAGCCAGACGAGGATGTGCTGGGGCCAG 3232  
QY 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
DB 3233 GGCGCGCGCGGCTCTGCGCTCCGAGGCGCTGCGAGTGTGTCGCCACCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
DB 3293 CTCAGCTGACTGACAGCCTGCTGACCTAGTCCCTCTCTGGGGTCACTCAGGAGCAGC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
DB 3353 CAGAGCAGCTGAGTCCGAGCTCCCGGGAGCAGCTGACTGCCCTGGAGGCGCGCAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
DB 3413 AACCGGCACTGCCCTCAGACTTCAAGACCATCTCGAC 3451

## RESULT 12

US-10-044-539-1  
Sequence 1, Application US/10044539  
Publication No. US20030100093A1  
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/044,539  
FILING DATE: 11-Jan-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002600US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4015 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 56..3454  
 OTHER INFORMATION: /product= "hTRT"  
 /note= "human telomerase reverse  
 transcriptase (hTRT) catalytic protein  
 component"  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

-10-044-539-1

Alignment Scores:  
 ed. No.: 0 Length: 4015  
 ore: 3505.00 Matches: 719  
 cent Similarity: 72.94% Conservative: 122  
 st Local Similarity: 62.36% Mismatches: 260  
 ery Match: 59.40% Indels: 52  
 : 15 Gaps: 13

-09-042-460-2 (1-1122) x US-10-044-539-1 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCCGCGCTCCCGCTCCGAGCGCGTCCCTGCTCGCAGCCACTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
 116 GTGCTGCGGTGCGCCAGTTCGTGCGCGCGCTCGGCGCCCGCGGCTGCGGTGCGAG 175  
 41 ProGlyAspProLysLeuValArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGCGGGGACCGCGCGCTTCCCGCGCTGCTGGCCAGTGGCTGCTGCTGCGCTCGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuLeu 80  
 236 GACGACGCGCGCGCGCGCTTCCCTGCGCGAGTGTCTCGCTGCTGCTGCTGCTGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGCCCCGAGTGTCTGAGAGGCTGTGCGAGCGCGCGCGCGAGAGTGTGCTTCCGCT 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGTGGACGGGGCGCGCGGGCGCGCGGGCGCGCGGGCGCGCGGGCGCGGG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaThrMetLeuLeu 140  
 416 AGCTACCTGCCCCAACCGGTGACCGACGACACTCGCGGGGAGCGGGGCGGTGGCTGCTG 475

QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 DB 476 CTGCCCGCGTGGCGGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 DB 536 CTGGTGGCTCCAGCTCGCTACGAGTGTGGGGCGCGCTGTACGAGTGTGGCGCT 595  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 DB 536 GCCACTCAGGCG 549  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220  
 DB 650 -----TGCAGCGCGCGCTGGAACCATAGCTCAGGAGGCGCGCG 688  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 DB 689 GTCCCCCTGGGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 DB 749 CTGCGGTGCGCCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 808  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 DB 809 GGGCAGGCGTCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 DB 869 GTGGGTGCTCCTGCGAG-----CCGCCGAGAGAGCCACCTCTTTGGAG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 914 GGTGGCTCTCTGGCAGCGCGCACTCCACCCATCTCGTGGCGCGCGCGCGCGCGCG 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 974 CCCCCATCCATCG 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGGAC---AAGGAGCAGCTGCGCGCTCTTC 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370  
 DB 1091 CTACTCAGCTCTTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 DB 1151 CTGGGTTCAGGCGCTGTGATCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1211 TACTGGCAATCGCGCGCGCTTCTGAGAGTGTGTGGAAACACCGCGCGAGTCCCGCTAC 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 DB 1271 GGGGTGCTCTCTCAGACGCACTGCGCGCTGCGAGCTGGG-----GTACCGCGCA 1321  
 QY 431 Leu-----Asn 432  
 DB 1322 GCCGGTGTCTGTGCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 DB 1382 ACAGACCGCGCGCGCTGTGTGAGCTGTCTCGCGCGCGCGCGCGCGCGCGCGCGCG 1441  
 QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 DB 1442 GGTCTGTCGCGCGCTGCTGCGCGGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501

473 AsnGluArgArgPhePheLeuLeuLysPhePheLeuLeuLysTyrGlyLys 492  
1502 AACGAACGCGCTTCTCAGGAACACCAAGAGTTCATCTCCTGCGGAAGCATCCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGCTGACGAGAGTACGCTGGAGATGAGCGTGGGAGTCTGGCTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLysLeuAla 532  
1622 AGCCACAGGGGTGGCTGTGTTCGGCGCGCAGACACCGCTCTGCGTGAGGAGATCCTGGCC 1681  
533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGCTGAGTGTGTGCTGTGAGCTGTCTAGGCTTCTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572  
1742 GTCACGGAGACACGCTTCAAAAGACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTGTCAAGACATTTGAATCAGACACACTTGAAGAGGTCGACGTGGGAGCTGTG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTyrLeuAlaMetPheProLysCysArgLeuArg 612  
1862 GAACGAGAGGTTCAGCAGCATTCGGAGACGAGCCGCTGTGTGAGCTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProLysValAsnMetSerTyrSerMetGlyThr 632  
1922 TTATCTCCCAAGCCTGACGGCTCGCGCCGATTTGTGAACATGAGTACGTCTGTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGACGTTCCAGAGAAAGAGGCGCGAGCTCTACCTCGAGGGTGAAGCAGCTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGCGCGCGCGCCGCTCTGTGGCGCTCTGTGTGGCG 2101  
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGAGGCTGTGGCGACCTCTGTGTGTGTGTGGCGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCTGTGAGCTGTACTTGTCAAGTGTGTGAGCGCGCGTACGACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCACGAGGTCATCGCAGCATCATCAAA---CCCGACAGACAGTACTCGTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 GTGCGTATGCGTGGTCCAGAGCGCGCCCGACGTCCGCGAGGCTTTCAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTTACTTGTACAGACCTCCAGCGCTACATCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGCTGTATCGACGAGCTCTCTCCCTGAA 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGACAGTGGCTCTTCGAGCTCTTCTACGCTTCTATGTGTGCCACGCGGTGCGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTCAGTCCAGTCCAGGGATCCCGCAGGGCTTCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852

Db 2573 CTGCTCTGAGCGCTGTGCTTACGCGACATGAGAACAGCTGTGTTCGGGGATTCGGCGG 2632  
Qy 853 AspGlyLeuLeuArgPheValAspPhePheLeuLeuValThrProHisLeuAspGln 872  
Db 2633 GACGGGCTGCTCTGCGTTTGGTGGATGATTTCTTGTGTGGTACACCTCACCTCACCCAC 2692  
Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
Db 2693 GCAGAACCTTCTCAGGACCTGTGTCGAGGTGCTTCCCTGAGTATGGCTGGTGGTGAAC 2752  
Qy 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
Db 2753 TTGCGGAGACAGTGTGTGACTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTT 2812  
Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuAspThrGlnThr 932  
Db 2813 GTTCAGATGCGCGCCACGCGCTATTCCCTGTGTCGCGCTGCTGCTGGATACCGGAGC 2872  
Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2873 CTGAGGTGACAGCGACTTCTCAGCTATGCCCCGAGCTTCATCGAGCGCAGCTCTAC 2932  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2933 TTCACCGCGGCTTCAAGGCTGGAGAACATGCTGCGAAACTCTTTGGGCTCTTGGG 2992  
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2993 CTGAAGTGTACACGCTGTCTTCTGATTTGCGAGTGAACAGCTCCAGCGGTGTGAC 3052  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGTCGAGCGGTACAGGTTTCAGCATGTGTGTGCGAGCTC 3112  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTCTGCGGCTCATCTCTGACAGC 3172  
Qy 1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCTCTCTCTGTACTCTCATCTCTGAAGCAGACAGCGAGGATGTCTGCGGGCCAG 3232  
Qy 1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCTCTGCGCTCCGAGCGGTGAGTGTGCTGTGCGCACCAAGCATTCCTG 3292  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGCTGACTCGACACCGCTGTACCTACGTGCGACCTCTCTGGGTCACTCAGGACGCC 3352  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGCGAGCTGAGTCCGAAGCTCCCGGAGCGAGCTGACTGTGCTGAGCGCGCAGC 3412  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGCGACTGCGCTCAGACTTCAAGACCATCTCGAC 3451  
RESULT 13  
US-10-385-882-1  
; Sequence 1, Application US/10385882  
; Publication No. US20030232409A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA  
; APPLICANT: FARRIS, James  
; APPLICANT: FOSTER, Douglas  
; APPLICANT: O'GRADY, Scott  
; TITLE OF INVENTION: IMMORTAL PORCINE CELLS  
; FILE REFERENCE: 110.01700101  
; CURRENT APPLICATION NUMBER: US/10/385,882  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: 60/363,129

PRIOR FILING DATE: 2002-03-11

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 4027

TYPE: DNA

ORGANISM: ARTIFICIAL

FEATURE:

OTHER INFORMATION: Nucleotide sequence encoding telomerase reverse transcriptase

-10-385-882-1

Alignment Scores:

Seq. No.: 0 Length: 4027

Score: 3498.00 Matches: 718

Percent Similarity: 72.85% Conservativity: 122

Percent Local Similarity: 62.27% Mismatches: 261

Gap Match: 59.28% Indels: 52

Gaps: 13

-09-042-460-2 (1-1122) x US-10-385-882-1 (1-4027)

```
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
57 ATGCGCGCGCTCCCGCGTCCGAGCGTGGCTCCCTGCTGGCGAGCCACTACCGGAG 116
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgLeuValGln 40
117 GTGCTCCCGCTGGCCACGTTCTGGCGGCGCTGGGCGCCCGCCAGGCTGGCGCTGGTGCAG 176
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
177 CGCGGGAGCCCGCGGCTTTCGCGCGTGTGTGGCCAGTGCCTGGTGTGGCTGCCCTGG 236
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80
237 GAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 296
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
297 GTGCGCGCGAGTGTGAGAGGCTGTGGAGCGCGCGCGCGCGAGACGCTGCTCGCTCGGC 356
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
357 TTGCGCGCTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 416
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
417 AGTACTGCTCCCAACACGTGACCGACGACCTGGCGGGGAGCGCGCGCGCTGGGGCTGCTG 476
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
477 TTGCGCGCGTGGCGGACGACGCTGTGTTACCTGCTGGCACGCTGGCGCTCTTTGTG 536
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
537 CTGTGTGCTCCAGCTGCGCTACACAGGTGTGGCGCGCGCGCGCGCGCTGTACAGCTGGCG 596
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
597 GGCACCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
651 -----TCGAAACCGGCGCTGGAAACCATAGCGTCAGGAGCGCGCGG 689
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
690 GTCCCGCTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 749
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro---- 259
750 CTGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 809
```

```
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
Db 810 GGGCAGGGGCTCTGGGCGCCACCGCGGACGCGTGGACCGAGTGCAGCGTGTTCGT 869
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
Db 870 GTGTGTACACTGCCAGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 914
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 915 GGTGGCTCTCTGGCGACGCGCCACTCCACCTCCGCTGGCGCGCGCCACCGCGCGGC 974
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 975 CCCCCATCCATCGCGCGCCACCGCTCCCTGGGACACGCTTGTCCCCGGGTGACGCC 1034
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1035 GAGACCAAGCACCTTCTCTACTCTCTCAGCGCAGC---AAGGAGCAGCTGGCGCTCTTC 1091
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
Db 1092 CTACTCAGCTCTCTGAGCGCCCGAGCTGACTGGCGCTCGGAGGCTCTGGAGACCATCTTT 1151
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1152 CTGGTTCCAGCGCTCTGATGCGCAGGACTCCCCGCGAGTTGCCCGCTGCCAGCGC 1211
QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
Db 1212 TACTTGGCAATGCGCGCGCTCTTCTGGAGCTCTTGGGAACACCGCGCAGTCCCTAC 1271
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
Db 1272 GGGTGTCTCTCAGACGCACTGCGCGCTGGAGCTGG-----GTCACCCCGACGA 1322
QY 431 Leu-----Asn 432
Db 1323 GCGCGTGTCTGTGCGCGGAGAACCCCGAGGCTCTGTGCGCGCGCGCGAGGAGGAC 1382
QY 433 ThrSerProProHisLeuMetAspLeuLeuLeuHisSerSerProTrpGlnValTyr 452
Db 1383 ACAGACCCCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1442
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
Db 1443 GGTCTGTGGCGCTGCTGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTC 1502
QY 473 AsnGluArgArgPhePheLysAsnLeuLysPheLysSerLeuGlyLysTyrGlyLys 492
Db 1503 AACGAAACCGCGCTTCTCAGGAAACACCAAGAGTTCTCTCTGGGAGAGCATGCCAAG 1562
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
Db 1563 CTCTCGCTCAGGAGCTGACGTGGAAGATGAGCTGGCGGCTGCGCTTGGCTGCGCAGG 1622
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluAlaLeuAla 532
Db 1623 AGCCGAGGGGTGGCTGTGTTCGCGCGCGAGACCGCTCTGGTGGAGAGATCCCTGGCC 1682
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552
Db 1683 AGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
Db 1743 GTACGAGGAGCACCTTTCAAAGAACAGGCTCTTTTTCACCGAAGAGTGTCTGAGC 1802
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
Db 1803 AGTTCGAAGAGCTTGGATCAGACGACCTTGAAGAGGTGAGCTGCGGAGCTGTCG 1862
QY 593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612
```

```
1863 GAAGCAGAGGTCAGCAGCATCGGGAAGCCAGGCGCCCTGCTGCTCCAGCATCCGC 1922
1864 .....
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1923 TTCAATCCCAAGCTTGAGCGGCTGCGCGCGATGTGAACTAGTACGTCGCGGAGCC 1982
633 ArgAlaLeuGlyArgArgGlyAlaAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1983 AGAACGTTCCGCGAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGGCATGTTT 2042
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
2043 ACGGTGCTCAACTACGAGCGCGCGCGCGCGCGCTCTCTGCGCGCTCTGCTGCTGGC 2102
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2103 CTGGAGCATATCCAGAGGCTTGGCGCACTTCTGCTGCTGCGGCGCCAGGACCGG 2162
693 ThrProArgMetTyrPheValAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2163 CGGCTGAGCTGACTTGTCAAGGTGATGTGACGGCGCGGTACGACACCATCCCGCAG 2222
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2223 GACAGGCTCACGAGGTATCGCCAGCATCATCAAA---CCCCAGAACAGTACTGCTG 2279
733 ArgGlnTyrAlaValAlaArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2280 COTCGGTATGCGGTGTCAGAGGCGCGCCCATGGCAGCTCCGCAAGGCTTCAAGAGC 2339
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2340 CACGTCTCTACCTTGACAGACTCCAGCGGTACATGCGACAGTTCGTGGCTCACTG 2399
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792
2400 GAGACC-----AGCCCGTGAAGGATCCCGTCTCATCGAGCAGAGCTCTCTCTGAT 2453
793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812
2454 GAGGCGCAGGTGGCTCTTGGAGCTCTTCTAGCGTCTTCTAGCGTCTTCTAGCGCTG 2513
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2514 ATCAGGGCGCAAGCTCTGCTACGTCAGTCCGCGGATCCCGAGGCTCCATCTCTCACG 2573
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2574 CTGCTCTGAGCGCTGCTACGCGGACATGAGAACAAAGCTGTTTGGCGGATTCGCGG 2633
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872
2634 GACGGGCTGCTCTCGCTTGTGGTGAATATTCTTGTGTGGTGACACCTCACTCACTCC 2693
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2694 GCGAAACCTTCTCAGACCTGTCAGAGGTGTCCTGAGTATGCTGCGGTGGTGAAC 2753
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2754 TTGCGGAAGACAGTGGTGAATCTTCCCTAGAACAGCAGGCGCTTGGTGGCGGCTTTT 2813
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2814 GTTCAGATGCGCGGCCACGGCTATTCCCTTGGTGGCGCTGCTGCTGATATCCCGGACC 2873
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2874 CTGGAGGTGCAGAGCGACTACTCTCAGCTATGCCCGACCTCCATCAGAGCGAGTCTACC 2933
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
```

```
Db 2934 TTCAACCGCGCTTCAAGGCTGCGGAGGAACATCGTGCAGAACTCTTTGGGGTCTTGGCG 2993
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
Db 2994 CTGAAGTGTCAAGCCCTGTTTCTGGATTTCAGGTGAACAGCTCCAGACGTTGTGCACC 3053
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
Db 3054 AACATCTACAAGATCTCTCTGCTGCTGAGGCGGTACAGGTTTCAGCATGTGTGTGCAGCTC 3113
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3114 CCATTTCATCAGCAAGTTTGGAAAGACCCACATTTTCTGCGCGTCTATCTTGACAGC 3173
Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
Db 3174 GCCTCCCTCTGCTACTCTCATCTCTGAAAGCCAGCAGGATGCTGCTGGGGGCCAAG 3233
Qy 1053 GlySer-----PheProGluAlaAlaHisTyrPleuCysTyrGlnAlaPheLeu 1069
Db 3234 GCGCCCGCGCGCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3293
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuGlyProLeuArgThrAla 1089
Db 3294 CTCAGCTGACTGCACACCGTGTACCTAGTGCCTCTCTGGGGTCACTCAGACAGCC 3353
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
Db 3354 CAGACGAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGCGCGCAGCC 3413
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3414 AACCGGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3452
```

## RESULT 14

```
US-10-205-629-1
; Sequence 1, Application US/10205629
; Publication No. US20030049236A1
; GENERAL INFORMATION:
; APPLICANT: Kassem, Moustapha
; APPLICANT: Jensen, Thomas
; APPLICANT: Rattan, Suresh
; TITLE OF INVENTION: Immortalized Stem Cells
; FILE REFERENCE: 006148.00002
; CURRENT APPLICATION NUMBER: US/10/205,629
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/315939
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PA 2001 01148
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-629-1

Alignment Scores:
Pred. No.: 0 Length: 3453
Score: 3496.00 Matches: 717
Percent Similarity: 72.85% Conservative: 123
Best Local Similarity: 62.19% Mismatches: 261
Query Match: 59.24% Indels: 52
DB: 15 Gaps: 13

US-09-042-460-2 (1-1122) x US-10-205-629-1 (1-3453)
```

```
Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20
Db 10 ATGCGGCGGCTCCCGCTGCGAGCGGTGCGCTCTCTGTCGCGCAGCCACTACCGCGAG 69
Qy 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
```

```

1105 CTGGTTCCAGGCGCTCGATGTCAGGAGCTCCCGCAGGTTGCCCGCGCTGCCCGCAGCGC 1164
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
1165 TACTGGCAATGCGGCCCTGTTTCTGGAGCTGTTGGGAACACACGCGCAGTGCCCTAC 1224
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
1225 GGGGTGTCCTCAAGACGCACTGCCCGCTCGAGCTGGC-----GTCACCCAGCA 1275
431 Leu-----Asn 432
1276 GCGGTGTCTGTGCCGCGAGAGCCCGCAGGCTCTGTGGCGGCCCGCAGGAGGAGGAC 1335
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
1336 ACAGACCCCGCTGCCCTGGTGAGCTGCTCCGACGACACACGAGGAGCCCTGGCAGGTGAC 1395
453 GlyPheLeuArgAlaCysLeuLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
1396 GGCCTCGTGGCGGCTGCTGCGCGGCTGTGCGCGCTGTGCCCGCAGGCTCTGGGGCTCCAGGCAC 1455
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492
1456 AACGAACCGCGCTTCTCAGGACACACCAAGATTCATCTCCCTGGGGAAGCATGCAAG 1515
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
1516 CTCCTGCTGCGAGGCTGACGTGGAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGG 1575
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgGlnLeuAla 532
1576 AGCCGAGGGTGGCTGTGTTCCGCGCGCAGACACCGCTCTGGCTGAGGAGATCCTGGCC 1635
533 ThrPheLeuPheTrpLeuMetAspThrTyrValGlnLeuLeuArgSerPheTyr 552
1636 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1695
553 IleThrGluSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrpSer 572
1696 GTCAGGAGACCGCTTCAAAAGAACAGGCTCTTTTCTACCGCGCGAGTGTCTGGAGC 1755
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1756 AAGTGTCAAGCATTTGGAATCAGACGACTTGAAGAGGTGAGCTGCGGAGCTGTGCG 1815
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612
1816 GAAGCAGAGGTTCAGCAGCATCGGAGCCAGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1875
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1876 TTCTATCCCAAGCTGAGCGGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1935
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1936 AGAAGCTTCCGAGAGAAAGAGGCGCGCTCTCCTCGAGGGTGAAGGCACCTGTTTC 1995
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
1996 AGCGTGTCTCAATACGAGCGCGCGCGCGCTCTGCGCGCTCTGCGCGCTCTGCTGCTGCTGCT 2055
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2056 CTGGACGATATCCACAGGCGCTGCGCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2115
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2116 CGCGCTGAGCTGACTTTGTCAAGTGTGATGTGAGCGCGCTGAGCAGACCATCCCGCAG 2175
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2176 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAAA-----CCCCAGAACACGTACTGCTG 2232
```





141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
4409 CTGCGCGCGTGGCGACGAGCTGTGGTTCACCTGTGGCAGCTCGCGCTCTTTGG 4468  
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
4469 CTGGTGGCTCCAGCTGGCTTACCAGGTGTGGCGCGCGCTGTACCGAGCTGGCGCT 4528  
181 ThrThrAspIleTrpProSerValSerTyrArgProThrArgProValGlyArg 200  
4529 GCCACTCAGCGCGCGCGCGCACACGCTAGT--GGACCCCGAAGCGCTCTGGGA-- 4582  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
4583 -----TGCAGCGCGCGCTGGAGACCTAACCTACGTCGAGGAGCGCGG 4621  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
4622 GTCCCTCTGGCGCTGCCAGCCCGCGGTGCGAGGAGCGCGCGGCGAGTGCACCGAAGT 4681  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluPro--- 259  
4682 CTGCGCTGCCAGAGGCCAGCGGTGGCGTGCCTGCCCTGAGCGGAGCGGCGCGCTT 4741  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
4742 GGGCAGGGGTCTGGCGCCACCGCGGAGGAGCGGTGACCGAGTGCCTGTCTCTGT 4801  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
4802 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGCACCTCTTTGGAG 4846  
293 GlyLysValSerAspLeuSerSer---GlySerValCysCysValHisLysProSer 311  
4847 GGTGCGCTCTCTGGCACCGCGCCACTCCACCCCTCCGTGGCGCGCCAGCACCGCGGGC 4906  
312 SerThrLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
4907 CCCCACCATCCGCGCGCACACGCTCCCTGGGACACGCTTGTCCCGCGGTACGCC 4966  
331 GluThrArgHisPheLeuTyrSerArgLysAspGlyGlnGluArgLeuAsnProSerPhe 350  
4967 GAGACAAGACACTTCTCTACTCTCCACGCGAC---AAGGAGCAGCTCGCGCCCTCTTC 5023  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
5024 CTACTCAGCTCTTGAGGCCAGCTGACTGCGCTCGAGGCTCGTGAGACCATCTTT 5083  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
5084 CTGGGTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCGCTCCCGCAGCGC 5143  
391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
5144 TACTGCAATAGCGGCCCTCTTCTGGAGCTGCTTGGACACGAGCGAGTGGCCCTAC 5203  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
5204 GGGGTGTCTCTCAAGACGCACTGCCGCTGCGAGTGGG-----GTCACCCCGACA 5254  
431 Leu-----Asn 432  
5255 GCGGTGTCTGTGCGGAGAGAGCCCGAGGCTCTGTGGCGCGCCCGAGGAGGAGC 5314  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
5315 ACAGACCCCGCTCGCTGTGTCAGCTGCTCCCGCAGCACAGCAGCGCCCTGGCAGGTGAC 5374  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
5375 GGCTTGTGCGGCGCTGCTGCGCGCGCTGGTGGTGCCTCCAGGCGCTCTGGGCTCCAGGAC 5434

473 AsnGluArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
5435 AACGAACCGCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCTGGGAGCATGCCAAG 5494  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
5495 CTCTCGCTGAGGAGCTGACGTGGAAGATGAGCGTGGGACTGGCGTGGCTGGCGCAG 5554  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
5555 AGCCACGCGGTGCTGTGTTCGCGCGCAGAGCACCGTCTCGTGAGGAGATCCTGGCC 5614  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
5615 AAGTTCCTGCTGCTGCTGATGAGTGTAGCTGCTGAGCTGCTCAGGCTCTTCTTTAT 5674  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
5675 GTACGAGGAGCACGTTTCAAAAGAACAGGCTCTTTTCTACGCGCGAGTGTCTGGAGC 5734  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
5735 AAGTTGCAAGCAATGGATTCACAGACGACTTGAAGAGGTGCAGCTGGGAGCTGTG 5794  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
5795 GAAGCAGAGGTTCAGGACGATCGGAAGACGAGCGCGCGCTGTGACGCTCCAGACTCCGC 5854  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
5855 TTATCCCGAAGCTGACGGCTGCGCGCATTTGGAATGACATGACATGCTGTGGAGGCC 5914  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
5915 AGAAGTCCCGCAGAGAAAGAGCGCGCGCTCACCTCGAGGCGTGAAGGCACTGTTC 5974  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
5975 AGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCGCGCGCTCTGTGCTGGCG 6034  
673 MetAsnAspIleTyrArgThrTrpAlaPheValLeuArgValArgAlaLeuAspGln 692  
6035 CTGACGATATCCACAGGCTGGCGCACCTTCTGCTGCTGCTGCGCGCCAGGACCCG 6094  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
6095 CCGCTGACCTGCTTGTCAAGTGGATGTACGCGCGCTGACGACCATCCCGCAG 6154  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
6155 GACAGGCTCAGGAGGTATCGCCAGCATCATCAA---CCCCAGAACAGTACTCGTG 6211  
733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
6212 CGTGGTATCGGTGCTCAGAGCGCGCCATGGGCACGTCGCGCAAGCGCTTCAAGAGC 6271  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
6272 CACGCTCTACTCTTGACAGACCTCAGCGGTACATGGCAGTCTGGGTCTACCTGCGAG 6331  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
6332 GAGACC-----AGCCGCTGAGGATCGCTGCTCATCGAGCAGAGCTCCTCCCTGAAT 6385  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValLys 812  
6386 GAGGACAGCAGTGGCTCTTCGACGCTTCTCAGCTTCATGTGCCACACCGCGCGTGGC 6445  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
6446 ATCAGGCGCAGTCTCTAGCTGCGAGGGATCCGCGAGGGCTCCATCTCTCCAG 6505  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

on: February 1, 2004, 15:05:59 ; Search time 5256 Seconds  
(without alignments)  
5188.284 Million cell updates/sec

le: US-09-042-460-2  
fect score: 5901  
pence: 1 NTRAPCPAVALSLRSRYRE.....TILKAADPALSTDFOTILD 1122

ring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

urched: 22781392 seqs, 12152238056 residues

al number of hits satisfying chosen parameters: 45562784

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

mand line parameters:

DEL=frame\_p2n.model DEV=xlh  
:/cgn2\_1/USPTO\_spool/US09042460/runat\_01022004\_150551\_17026/app\_query.fasta\_1.1287  
i=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
ITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
CALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
JFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
ER=US09042460 @CGN 1.1 3909 @runat\_01022004\_150551\_17026 -NCPU=6 -ICPU=3  
V\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -DSBLOCK=100 -LONGLOG  
APEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

abase :  
1: em estba.\*  
2: em esthum.\*  
3: em estin.\*  
4: em estmu.\*  
5: em estov.\*  
6: em estpl.\*  
7: em estro.\*  
8: em htc.\*  
9: gb estl.\*  
10: gb est2.\*  
11: gb htc.\*  
12: gb est3.\*  
13: gb est4.\*  
14: gb est5.\*  
15: em estfun.\*  
16: em estom.\*  
17: em ges hum.\*  
18: em ges inv.\*  
19: em ges pln.\*  
20: em ges vrt.\*  
21: em ges fun.\*  
22: em ges mam.\*  
23: em ges mus.\*  
24: em ges pro.\*  
25: em ges rod.\*  
26: em ges phg.\*  
27: em ges vrl.\*  
28: gb gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1455	24.7	851	13	BU702370	BU702370 UI-M-F10-
2	1279.5	21.7	851	13	BG917907	BG917907 602820830
3	1173	19.9	664	13	BQ258274	BQ258274 NISC kp11
4	1082	18.3	641	28	AZ972318	AZ972318 2M0246F07
5	938.5	15.9	925	12	BM453198	BM453198 AGENCOURT
6	841	14.3	599	10	BB618671	BB618671 BB618671
7	783	13.3	866	10	BE371943	BE371943 601217728
8	768.5	13.0	614	10	BB651920	BB651920 BB651920
c	595	10.1	347	9	AW244516	AW244516 BR END06B
10	558	9.5	422	12	BM824748	BM824748 K-SST0096
11	540.5	9.2	753	13	BU452535	BU452535 603767927
12	511	8.7	326	13	BY149368	BY149368 BY149368
13	493.5	8.4	775	12	BI388013	BI388013 BF126_002
14	483	8.2	880	13	BU377259	BU377259 603811228
15	463	7.8	389	9	AA281296	AA281296 zt08G02.r
16	462	7.8	668	14	CA380121	CA380121 659344 NC
17	435	7.4	696	13	BU139751	BU139751 603134527
18	427.5	7.2	632	14	CA353864	CA353864 625469 NC
19	407	6.9	568	4	BX521269	BX521269 RZPD Mus
20	405	6.9	534	9	AW318894	AW318894 un09a02.y
21	396.5	6.7	835	13	BU111946	BU111946 603127372
22	392.5	6.7	875	13	BU122597	BU122597 603148441
23	337	5.7	715	10	BE396925	BE396925 601290610
24	335	5.7	610	10	BE514188	BE514188 601316376
25	328	5.6	649	10	BE514070	BE514070 601316376
26	327	5.5	409	9	AA311750	AA311750 EST182469
27	325.5	5.5	813	13	BU224024	BU224024 603798349
28	321	5.4	679	10	BE396606	BE396606 601289077
29	293	5.0	779	10	BE268183	BE268183 601125261
30	282.5	4.8	1455	29	CC190951	CC190951 CH261-380
31	279.5	4.7	1230	29	CC211572	CC211572 CH261-14F
32	279.5	4.7	1554	29	CC190875	CC190875 CH261-38M
33	272	4.6	554	28	AQ397020	AQ397020 mgx00014C
c	266	4.5	1020	29	CNS05F1M	AL334435 Tetraodon
35	250.5	4.2	570	12	BM521744	BM521744 sak70f07
36	244.5	4.1	502	10	BE251764	BE251764 EST419026
37	216.5	3.7	487	28	AZ799615	AZ799615 2M0057L07
38	214.5	3.6	739	13	EX315053	EX315053 EX315053
c	212	3.6	696	28	BH962910	BH962910 cdj2ib08
40	209.5	3.6	393	13	BX088059	BX088059 BX088059
41	205.5	3.5	393	10	BF251775	BF251775 EST419037
c	203.5	3.4	625	28	B27802	B27802 T1704TRD TA
43	198	3.4	613	10	BF597086	BF597086 su83g03.y
44	197	3.3	1009	29	CNS0730U	AL427780 clone BA0
c	167.5	2.8	724	13	EX315052	EX315052 EX315052

# ALIGNMENTS

RESULT 1  
BU702370  
LOCUS  
DEFINITION  
IMAGE: 6400523 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

BU702370  
UI-M-F10-byx-f-12-0-UI.r1 NIH ENMAP\_F10 Mus musculus cDNA clone  
IMAGE: 6400523 5', mRNA sequence.  
BU702370  
BU702370.1 GI:23627105  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 851)

# AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

## FEATURES

source

1..851  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6400523"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 12.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP\_F10"  
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCCACGAC. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator." 220 t 3 others

SE COUNT 180 a 227 c 221 g 220 t

IGIN

ignment Scores:

ed No.: 1..17e-118 Length: 851  
ore: 1455.00 Matches: 278  
cent Similarity: 98.58% Conservative: 0  
st Local Similarity: 98.58% Mismatches: 4  
ery Match: 24.66% Indels: 0  
Gaps: 0

-09-042-460-2 (1-1122) x BU702370 (1-851)

729 ThrTyrCysIleArgInTyrrAlaValAlaArgAspSerGlnGlnValHislys 748  
6 ACGTACTGTATCCGACAGTATGAGTGGTCCGAGAGATAGCCAGGCCAAGTCCACAG 65  
749 SerPheArgGlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeu 768  
66 TCCTTTAGGAGACAGTCCACCCCTCTGTGACCTCCAGCCATACATGGCCAGTTCCT 125  
769 LysHisLeuGlnAspSerAspAlaSerAlaLeuArgAsnSerValIleGlnSer 789  
126 AAGCATCTGCAGGATTTCAGATGCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 185  
789 IleSerMetAsnGlnSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHis 808  
186 ATCTCTGAATGAGACACACACCCCTTTGACTTCTCTGCTGCTTCTCTGCTGCTC 245  
809 SerValValIleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySer 828  
246 AGTGTCTGAAGATTGGTGACAGGTGCTATACGACGTGCGAGGCATCCCCAGGCCTCC 305

QY 829 SerLeuSerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAla 848  
Db 306 AGCCTATCCACCCCTGCTCTGCACTGTGTTTCGAGAGACATGAGAACCAAGCTTTGCT 365  
QY 849 GluValGlnArgAspGlyLeuLeuArgPheValAspPheLeuLeuValThrPro 868  
Db 366 GAGTGTGAGCGGATGGGTGCTTTTACGTTTGTGAGTCTTCTGTTGTGAGGCT 425  
QY 869 HisLeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGly 888  
Db 426 CACTTGGACCAAGCAAAACCTTCTCAGCACCCCTGCTCCTGAGTATGGG 485  
QY 889 CysMetIleAsnLeuGlnLysThrValValAsnPheProValGlnProGlyThrLeuGly 908  
Db 486 TGCATGATAAACTTGCAGAGACAGTGGTGAACCTCCCTGTGAGCCTGTACCTGGT 545  
QY 909 GlyAlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeu 928  
Db 546 GTGTCAGCTCCATACACAGCTGCTGCTCACTGCTGCTTTCCCTGGTGTGGCTTGTGCTG 605  
QY 929 AspThrGlnThrLeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLys 948  
Db 606 GACACTCAGACTTTGGAGGTGTTCTGTGACTACTCTAGGTTATGCCAGACCTCAATTAG 665  
QY 949 ThrSerLeuThrPheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeu 968  
Db 666 ACGAGCTCCTCTCCAGAGTCTTTCANAGCTGGGAGACCATGGGAAACANGCTCCTG 725  
QY 969 SerValLeuArgLeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGln 988  
Db 726 TCGGTCTTGGGTGGAAGTGTCAAGTGTCAAGTGTATTTCTAGACTTCGAGGTGACAGCTCCAG 785  
QY 989 ThrValCysIleLeuIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCys 1008  
Db 786 ACAGTCTGCATCAATATATACAGATCTTCTGCTTCANGGCTACAGGTTCATGTCATGT 845  
QY 1009 Vallile 1010  
Db 846 GTGATT 851

RESULT 2  
BG917907

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 851)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10903 row: k column: 08

High quality sequence stop: 753.

Location/Qualifiers

1..851

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

FEATURES

source

/db xref="taxon:10090"  
 /clone="IMAGE:4949887"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Mam6"

/note="Organ: mammary; Vector: pCMV-SPORE6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally; Primer: Oligo dt.  
 Library constructed by Life Technologies; Investigator  
 providing samples: Jeffrey Green, M.D., NIH"

E COUNT 184 a 238 c 214 g 215 t  
 GIN

gment Scores:  
 d. No.: 4, 71e-103 Length: 851  
 re: 1273.50 Matches: 258  
 cent Similarity: 95.93% Conservative: 1  
 t Local Similarity: 95.56% Mismatches: 10  
 ry Match: 21.68% Indels: 4  
 12 Gaps: 0

09-042-460-2 (1-1122) x BG917907 (1-851)

850 ValGlnArgAspGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuValThrProHis 869  
 13 GTGCAGCGGGATGGGTG-CATTTACGTG-TGTGATGACTTTCTGTAGTGACGGCTCAC 70  
 870 LeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCys 889  
 71 TTGACCAAGCAAAACCTTCTTCAGACCCCTGGTCTCATGGCGTTCTGAGTATGGGTGC 130  
 890 MetIleAsnLeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGly 909  
 131 ATGATAAATTGCAGAGACAGTGGTGAATCTCCCTGTGGAGCTGGTACCTGGGTGGT 190  
 910 AlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAsp 929  
 191 GCAGTCCATACCAAGCTGCTGCTCACTGCTGTTTCCCTGGTGGCTGTGCTGGAC 250  
 930 ThrGlnThrLeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThr 949  
 251 ACTCAGACTCTGGAGGTGTTCTGTGACTACTCAGGTATATGCCAGACCTCAATTAGAGC 310  
 950 SerLeuThrPheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSer 969  
 311 AGCTCACTTCCAGAGTGTCTTCAAGCTGGGAAGACCACTGCGGAACAAGCTCCTGTGCG 370  
 970 ValLeuArgLeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThr 989  
 371 GTCTTGGGTGAAGTGTACGGTCTATTTCTAGACTTGCAGGTGAACAGCTCCAGACA 430  
 990 ValCysIleAsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysVal 1009  
 431 GTCCTCATCAATATATACAGATCTTCTCTCTTCAGGCTTACAGGTTCATGATGATGTG 490  
 1010 IleGlnLeuProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIle 1029  
 491 ATTCAAGCTTCTCTTGTACCAAGCGGTGTAGAAAGACCTCATCTCTTCTGGGCATCATC 550  
 1030 SerSerGlnAlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeu 1049  
 551 TCCAGCCAAGCATCTCTGCTGTATCTCTCTCTTCAAGGTCAAGATTCAGGAATGACACTA 610  
 1050 LysAlaSerGlySerPheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
 611 AAGGCCCTGTGCT 669  
 1070 -LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAl 1089  
 670 GCTCAAGCTGGTCT 729  
 1089 aGlnLysLeuLeuCysArgLysLeuProGluAlaIleThrMetThrIleLeuLysAlaAla 1109

Db 730 CCAAAACCGCTGTCCGGAAGCTTCGAGAGCGGACATGACCATGCTTACAGTGCAGC 789  
 Qy 1109 aAspProAlaLeuSerThrAspPheGln 1118  
 Db 790 TGACCCAGACCTAAGGACAGACTTTTCAG 817

RESULT 3  
 BQ258274 664 bp mRNA linear EST 06-MAY-2002  
 LOCUS NISC kp1lg04.q3 Baker mouse embryo e7.5 Mus musculus cDNA clone  
 DEFINITION IMAGE:5409222, mRNA sequence.

ACCESSION BQ258274  
 VERSION BQ258274.1 GI:20459030  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 664)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 cDNA Library Preparation: J. Baker (Stanford University)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 MGI:1845958

Plate: LLAM12043 row: N column: 7  
 Seq primer: Sp6 primer.  
 Location/Qualifiers  
 1..664

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5409222"  
 /tissue\_type="embryo, late gastrula"  
 /dev\_stage="embryo, 7.5 dpc"  
 /lab\_host="XLI-Blue"  
 /clone\_lib="Baker mouse embryo e7.5"  
 /note="Vector: pCG105; Site 1: NotI; Site 2: SalI; cDNA  
 made by oligo-dT priming. Directionally cloned into  
 SalI/NotI sites using the following 5' adaptor:  
 5'-TCGACCCAGCGCTCG-3'. Size-selected for average insert  
 size 1.8-1.9 kb. Library constructed by J. Baker (Stanford  
 University)."  
 BASE COUNT 155 a 177 c 183 g 149 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9,31e-94 Length: 664  
 Score: 1173.00 Matches: 218  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.54% Mismatches: 0  
 Query Match: 19.88% Indels: 0  
 DB: 13 Gaps: 0

US-09-042-460-2 (1-1122) x BQ258274 (1-664)

Qy 356 GlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePheLeuGlySerArgPro 375  
 Db 7 GAGCCTAACTTGACTGGGCGCCAGGAGCTGGTGGAGATCATCTTCTGGGCTCAAGGCT 66  
 Qy 376 ArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArgTyrTyrGlnMetArg 395  
 Db 67 AGGACATCAGGACCACTCTGCGAGGACACACCGCTCTATCGCGTGGATCTGCGCAGATCGG 126



```

396 ProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyrValArgLeuLeuArg 415
127 CCCCTGTTCCACAGCTCTCGTGAACATCCAGAGTCCCAATATGTCAGACTCCCTCAGG 186
416 SerHisCysArgPheArgThraAlaGlnGlnValThrAspAlaLeuAenThrSerPro 435
187 TCACATTGCAGGTTTCGACAGCAGCAACCAACAGTGCAGATGCTTGACACAGCCCA 246
436 ProHisLeuMetAspLeuArgLeuHisSerSerProTyrGlnValTyrGlyPheLeu 455
247 CGCACCTCATGATTGCTCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCTT 306
456 ArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHisAsnGluArg 475
307 CGGGCTGCTCTCTGCAAGGTGTGTCTGCTAGTCTCTGGGGTACAGGCACATGAGCG 366
476 ArgPhePheLysAsnLeuLysPheLysSerLeuGlyLysTyrGlyLysLeuSerLeu 495
367 CGCTTCTTTAAGACTTAAGAAAGTTTCATCTCGTTGGGMAATACGGCAAGCTATCACTG 426
496 GlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSerSerProGly 515
427 CAGGAACTCATGTGGAAGATGAAGATGAGGATTCCTGCTCCGCGAGGAGCCAGGG 486
516 LysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlnLeuAlaThrPheLeu 535
487 AAGGACCGTGTCCCGCTCCAGAGCAGCCTGTGAGGAGAGAGTCTGCTAGTTCCTG 546
536 PheTyrLeuMetAspThrTyrValValGlnLeuArgSerPhePheTyrIleThrGlu 555
547 TTCTGCTGATGCACATACATCGTGTACAGCTGCTTAGGTCATTCTTTTACATCACAGAG 606
556 SerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSerLysLeu 574
607 AGCACATTCAGAGACAGACAGCTCTCTTCTTACCGTAAAGATGTGTGGAGCAGCTG 663

```

## SULT 4

```

972318
CUS 2M0246F07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
FINITION clone UUGC2M0246F07 F, genomic survey sequence.

```

```

SESSION AZ972318.1 GI:13843545
RSION AZ972318.1
YWORDS GSS.
ORCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 641)
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

```

```

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

```

```

JOURNAL Unpublished
MMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

```

```

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0246 row: F column: 07
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: Plasmid ends
High quality sequence stop: 641.
Location/Qualifiers
1. .641
/organism="Mus musculus"

```

## ATURES

```

source

```

```

/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0246F07"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

```

BASE COUNT 155 a 195 c 157 g 134 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.09e-85 Length: 641
Score: 1082.00 Matches: 203
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 18.34% Indels: 0
DB: 28 Gaps: 0

```

```

US-09-042-460-2 (1-1122) x AZ972318 (1-641)

```

```

QY 244 AlAlsLysAlaArgCysTyrProValProArgValGluGluGlyProHisArgGlnVal 263
Db 26 TCGCAGAAGGCCAGATCTATCTCTCCGAGAGTGGAGGAGGCCACACAGGCGAGTG 85
QY 264 LeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerProGlnValPro 283
Db 86 CTACCAACCCCATCAGCAAAATCATGGTGCCAAAGTCTCGTCCGTCCTCCGAGGTCCT 145
QY 284 ThrAlaGluLysAspLeuSerSerLysGlyLysValSerAspLeuSerLeuSerGlySer 303
Db 146 ACTGCAGNAGAAAGATTGTCTTCTTAAGGAAGAGTGTCTGACCTGAGTCTCTCTGGTGC 205
QY 304 ValCysCysLysHisLysProSerSerThrSerLeuLeuSerProProArgGlnAsnAla 323
Db 206 GTGTGCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACCCCGCCAAATGCC 265
QY 324 PheGlnLeuArgProPheIleGluThrArgHisPheLeuTyrSerArgGlyAspGlyGln 343
Db 266 TTTGAGTCTCAGGCCATTTATGAGACCCAGACATTTCTTTACTTCCAGGGAGATGGCCAA 325
QY 344 GluArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArg 363
Db 326 GAGCGTCTAAACCCCTCATTTCTTACTCTCAGCACTCAGCCCTAACTTGTAGTGGGCGAG 385
QY 364 ArgLeuValGluIleIlePheLeuGlySerArgProArgThrSerGlyProLeuCysArg 383
Db 386 AGACTGGTGGAGATCATCTTTCTGGGGCTCAAGGCCATGAGACATCAGGACCATCTCAGG 445
QY 384 ThrHisArgLeuSerArgArgTyrTyrGlnMetArgProLeuPheGlnGlnLeuLeuVal 403
Db 446 ACACACCGCTCATCGCTCGATCTGCGAGATCGGGCCCTGTGTCCACAGCTGCTGGTG 505
QY 404 AsnHisAlaGluCysGlnTyrValArgLeuLeuArgSerHisCysArgPheArgThrAla 423

```

506 AACATGACAGTGCCTATATGTCAGACTCCTCAGGTACAGTTCAGAGTTTCACACACA 565  
 424 AsnGlnValThrAspAlaLeuAnThrSerProHisLeuMetAspLeuArg 443  
 566 AACACAGGTGACAGTGCCTTGAACACAGCCACCGCCTCATGGATTGCTCGC 625  
 444 LeuHisSerPro 448  
 626 CTGCACAGCAGTCCC 640

BLAST 5  
 US  
 INITIATION 6387556 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5529840  
 5', mRNA sequence.  
 BSSION BM453198  
 WORDS BM453198.1 GI:18502238  
 RCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 925)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM2208 row: p column: 01  
 High quality sequence stop: 646.  
 Location/Qualifiers  
 1. 925  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5529840"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."  
 172 a 290 c 276 g 187 t

TURES  
 source

gment Scores:  
 d. No.: 1.13e-72 Length: 925  
 re: 938.50 Matches: 186  
 cent Similarity: 79.64% Conservative: 37  
 t Local Similarity: 66.43% Mismatches: 52  
 ry Match: 15.90% Indels: 5  
 12 Gaps: 2

09-042-460-2 (1-1122) x BM453198 (1-925)

689 AlaLeuAspGlnThrProArgMetTyrPheValValAlaAspValThrGlyAlaTyrAsp 708  
 3 GCCCAGGACCGCGCTGAGCTGTACTTGTTCAGGTGGATGTGACGGCGCGTACGAC 62  
 709 AlaIleProGlnGlyLeuValGluValValAlaAsnMetIleArgHisSerGluSer 728  
 63 ACCATTCCCGAGGACAGCTCATCGGAGTTCATCGCAGCATCATCAA---CCCCAGAAC 119  
 729 ThrTyrCysIleArgGlnTyrAlaValValAlaArgAspSerGlnGlyGlnValHisLys 748  
 120 ACGTACTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 179

QY 749 SerPheArgArgGlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeu 768  
 Db 180 GCGTTCAAGACGACCGTCTCTACCTTGACAGCTCCAGCGGTATCGACAGCTTCGTG 239  
 QY 769 LysHisLeuGlnAspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSer 788  
 Db 240 GCTCACCTGCAGAGAC---AGCCCGCTGAGGGATGCGCTCATCGACAGAGC 293  
 QY 789 IleSerMetAsnGluSerSerSerLeuPheAspPhePheHisPheLeuArgHis 808  
 Db 294 TCCTCCTCATAGAGCGCAGCAGTGGCTCTTCGACGCTCTTCCTACGCTTCATGTGCCAC 353  
 QY 809 SerValValValIleLeuAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySer 828  
 Db 354 CACCGCGTGGCATCAGGGGCAAGTCTTCGTCAGTCCAGGGGATCCCGCAGGGTCC 413  
 QY 829 SerLeuSerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAla 848  
 Db 414 ATCTCTCCACGCTGCTCTGCAGCCTGTGTACGCGCATCGAGACAACAGCTGTTGGC 473  
 QY 849 GluValGlnArgAspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrPro 868  
 Db 474 GGGATTGCGCGGACGGGCTGCTCTCTGCTTGGTGGATGATTCTTGTGGTACACCT 533  
 QY 869 HisLeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGly 888  
 Db 534 CACCTCACCCACGCGAAGACCTTCTCAGGACCTGTCTCCGAGGTGCTCCTGAGTATGGC 593  
 QY 889 CysMetIleAsnGlnLysThrValValAsnPheProValGluProGlyThrLeuGly 908  
 Db 594 TGGGTGTGAATTCGGAAGACAGTGTGAATTCCTGTAGAGACGAGCGCTGGT 653  
 QY 909 GlyAlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeu 928  
 Db 654 GGCACGGCTTTGTTTCAGATGCCGCGCCACCGCTATTCCTGTTGGTGGCTGCTGCTG 713  
 QY 929 AspThrGlnThrLeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLys 948  
 Db 714 GATACCGGACCTCGAGGTGCAGAGCGACTACTCCAGCTATGCCGAGACCTCCATCAGA 773  
 QY 949 ThrSerLeuThrPheGlnSerValPheLysAlaGlyLys-ThrMetArgAsnLys 966  
 Db 774 GCAGTCTCACCTTCAACCGCGCTTCAAGGCTGGGAAGAACATGCGTCGCAAA 829

RESULT 6  
 BB618671 599 bp mRNA linear EST 26-OCT-2001  
 BB618671 RIKEN full-length enriched, 8 days embryo Mus musculus  
 LOCUS CDNA clone 5730412M20 5', mRNA sequence.  
 DEFINITION BB618671  
 BB618671  
 EST.  
 BB618671.1 GI:16458173

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 599)

REFERENCE  
 AUTHORS  
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,  
 M., Koya,S., Matsuura,T., Miyazaki,A., Nomura,K., Ohno,M.,  
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,  
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,I.,  
 Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

TITLE  
 JOURNAL  
 COMMENT  
 Unpublished  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222

Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10): 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11): 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2): 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

## UTURES

## Location/Qualifiers

1. .599  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="5730412M20"  
 /sex="mixed"  
 /dev\_stage="8 days embryo"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 8 days embryo"  
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was  
 prepared by using trisilose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of subtraction to  
 Ret = 100.0 Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGATTCGAGTTAATTAATTCACCCCCCCCCC 3']. cDNA  
 was cloned into the XhoI and BamHI sites. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC i. Cloning sites, 5' end: Sali; 3' end: BamHI."  
 SE COUNT 91 a 185 c 161 g 142 t

## IGIN

## Alignment Scores:

ed. No.: 2,61e-64 Length: 599  
 ore: 841.00 Matches: 161  
 cent Similarity: 100.00% Conservative: 0  
 st Local Similarity: 100.00% Mismatches: 0  
 ery Match: 14.25% Indels: 0  
 : 10 Gaps: 0

-09-042-460-2 (1-1122) x BB618671 (1-599)

1 MetThArgAlaProHrgCysProAlaValArgSerLeuLeuHrgSerArgTyrArgGlu 20  
 115 ATGACCCGCGCTCTCTGTTGCCCGCGGCGCTCTCTGCTGCGAGCGGATACCGGGAG 174

Qy 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgLeuValGln 40  
 Db 175 GTGTGGCCGCTGGCAACCTTTGTGGGCGCTGGGGCCGAGGCGGCGCTGTGCA 234  
 Qy 41 ProGlyAspProGlySileTyrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60  
 Db 235 CCGGGGAGCCCGAAGATCTACCGCACTTTGGTTGGTGGCCCAATGCTAGTGTGCTGCTG 294  
 Qy 61 GlySerClnProProAlaLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80  
 Db 295 GGCTCACAGGCTCCACCTCCGACCTTTCTTCCACAGGTGTCTATCCCTGAAGAGCTG 354  
 Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 355 GTGGCAGGGTGTGTCAGAGACTCTGCGAGCGCAACGAGAGAAAGCTGTGCTTTTGGC 414  
 Qy 101 PheGluLeuLeuHsngluAlaargGlyGlyProMetAlaPheThrSerSerValArg 120  
 Db 415 TTTGAGCTGTTTAAACGAGCGCAGAGCGGGCTCCCATGGCTTCACTAGTAGCGTGGT 474  
 Qy 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 Db 475 AGCTACTTGGCCCAACTGTATTGAGACCCCTGGTGCAGTGGTGCATGGATGCTACTG 534  
 Qy 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 535 TTGAGCGAGTGGCGAGCAGCTGCTGCTGCTACCTGCTGGCACACTGTGCTTATCTT 594  
 Qy 161 Leu 161  
 Db 595 CTG 597

## RESULT 7

BE371943

LOCUS

DEFINITION

601217728F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:3586614 5',

mRNA sequence.

ACCESSION

BE371943

VERSION

BE371943.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (Bases 1 to 866)

AUTHORS

NH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM8748 row: h column: 07

High quality sequence stop: 639.

Location/Qualifiers

1. .866

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:3586614"

/tissue\_type="spontaneous tumor, metastatic to mammary.

Stem cell Origin."

/lab host="DH10B"

/clone lib="NCI CGAP Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: Sali;

Site 2: NotI; Cloned unidirectionally; Primer: oligo dT.

Library constructed by Life Technologies. Investigator

## FEATURES

source



```

41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
235 CCGGGGACCCGAGGATCTACCGCATTTGGTTGCCAATGCCCTAGTGTGCATGCACCTGG 294

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
295 GGTCTCAGAGCTTCCACCTGCGACCTTCTCTCCACCGAGGTGTCACTCCCTGGAAGAGCTG 354

81 ValAlaArgValValGlnArgLeuCysLysArgGlnGluArgGlnValLeuAlaPheGly 100
355 GGGGACAGGTTGTGACAGACTCTGCGAGGCCAACGAGAGAAACGCTGCTGCTTTGGC 414

101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
415 TTGAGCTGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTTCTCATAGTACGGTGGCT 474

121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
475 AGCTACTTGGCCACACACTGTTATTGAGACCCCTGGCTGTGTCATGGTGCATGGTACTG 534

141 LeuSerArgVal-GlyAspAspLeuLeu---ValTyrLeuLeuAlaHisCysAlaLeuTy 159
535 TTGAGCCGAGTGGGCGGACGACCTGCTGGGTCTAACTGCTGGGACACACTGGGGCTCCTT 594

159 rLeuLeuVal 162
595 TATCCTCTCTG 604

```

```

347 bp mRNA linear EST 25-JAN-2002
BR END06B09 Bain Rancourt retinoic acid induced ES cell neural
differentiation subtraction library Mus musculus cDNA clone 06B09
similar to gb|AF073311|AF073311 Mus musculus telomerase catalytic
subunit mRNA, complete cds, mRNA sequence.

```

```

AW244516
AW244516.1 GI:8051265
EST.
Mus musculus (house mouse)
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bain,G., Mansergh,F.C., Wride,M.A., Hance,J.E., Isogawa,A.,
Rancourt,S.L., Ray,W.J., Yoshimura,Y., Tsuzuki,T., Gottlieb,D.I.
and Rancourt,D.E.
ES cell neural differentiation reveals a substantial number of
novel ESTs
Funct. Integr. Genomics 1 (2), 127-139 (2000)
21652683
11793228
Contact: Rancourt DE
Department of Biochemistry and Molecular Biology
University of Calgary
3330 Hospital Drive N.W., Calgary, Alberta, T2N 4N1, Canada
Tel: 403 220 2888
Fax: 403 283 8727
Email: rancourt@calgary.ca; URL:http://www.acs.ucalgary.ca/
rancourt
DNA sequencing by: University Core DNA Services, University of
Calgary. Submitted sequence has been trimmed at both ends to remove
the adaptor oligos containing the EcoRI sites: i.e. GAATTCGACTA
(beginning) and TAGTCGGAATTC (end) removed. Therefore, reported
insert length is longer than actual EST sequence length.
Insert length: 359 Std Error: 10.00
Seq primer: T3 Or T7.

```

```

1. .347
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129"
/db_xref="taxon:10090"

```

```

FEATURES
source

```

```

/clone="06B09"
/cell_type="embryonic stem (ES) cell"
/cell_line="D3"
/lab_host="DH5 alpha"
/clone_lib="Bain Rancourt retinoic acid induced ES cell
neural differentiation subtraction library"
/notes:Vector: pBluescript II SK+ (Stratagene); Site:1:
EcoRI; Site:2: EcoRI; Library constructed by Dr. Gerard
Bain (present address: Hoechst-ARIAD Genomic Center,
ARIAD Pharmaceuticals Inc., 26 Landsdowne Street,
Cambridge, Massachusetts, 02139-4234, U.S.A.). To isolate
cDNAs corresponding to mRNAs which are upregulated during
the neural differentiation of ES cells in vitro, the
subtractive hybridization technique of Wang and Brown [1]
was employed. Poly(A)+ RNA was prepared from both
undifferentiated ES cells and from embryoid bodies which
had been cultured for 4 days in the absence of RA followed
by an additional 3 days in the presence of 0.5 (M RA
(4-/3+ cells). These poly(A)+ RNAs were converted to
double-stranded cDNA using the Superscript Choice System
(Gibco). Aliquots of both cDNAs were digested with the
restriction enzymes Alul and Alul plus Real. An adaptor
oligo [1] containing an EcoRI site was ligated to the ends
of the restricted cDNAs to provide primer binding sites
and large amounts of each cDNA population were then
produced by the polymerase chain reaction (PCR) as
described [1]. Amplified cDNA from undifferentiated ES
cells was biotinylated using Photoprobe biotin (Vector
Laboratories) according to the manufacturer's protocol.
2.5 ug of amplified cDNA from 4-/3+ cells was mixed with
50 ug of biotinylated ES cell cDNA, denatured by boiling,
and hybridized for 20 h. Double stranded cDNAs containing
biotin were removed by streptavidin/phenol treatment as
described [1]. The remaining subtracted cDNA was mixed
with an additional 25 mg of biotinylated ES cell cDNA,
denatured by boiling, and hybridized for 2 h. The
streptavidin/phenol treatment was repeated and the
remaining cDNA was amplified by PCR [Wang and Brown, 1991
]. Two additional rounds of subtraction were repeated
exactly as described above. The cDNA obtained from this
subtraction procedure was digested with EcoRI and ligated
to pBS II SK+ (Stratagene) followed by transformation
into E. coli DH5 cells. Individual colonies were picked
and the corresponding plasmids were isolated either by an
alkaline lysis miniprep procedure [2], or using the
Qiaprep spin miniprep kit (Qiagen). Sequence analysis was
performed using the Big Dye Cycle Sequencing kit and an
ABI373 sequencer (University Core DNA Services,
University of Calgary). 1. Wang, Z; Brown, DD (1991) A gene
expression screen. Proc. Natl. Acad. Sc i. USA 88,
11505-11509. 2. Sambrook, J; Fritsch, EF; Maniatis, T. (1989)
Molecular Cloning: A Laboratory Manual. Cold Spring Harbor
Laboratory Press, Cold Spring Harbor, New York."

```

```

BASE COUNT      78 a      88 c      100 g      81 t
ORIGIN
Alignment Scores:
Pred. NO.:      Length:      347
Score:          595.00      Matches:      115
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.08%      Indels:      0
DB:              9          Gaps:      0

```

```

US-09-042-460-2 (1-1122) x AW244516 (1-347)

```

```

QY 731 CysIleArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPhe 750
Db 345 TGTATCCGCCAGTATGCAGTGTGTCGGAGAGATAGCCAGGCCAAGTCCACAGTCTCTT 286
QY 751 ArgArgGlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHis 770
Db 285 AGGACAGAGGTACACACCTCTCTGACCTCCAGCCATACATGCGCCAGTTCCTTAGCAT 226

```

771 LeuGlnAspSerAspAlaSerAlaLeuArgAsnSerValValleuGlnSerIleSer 790  
 225 CTGAGGATTCAGATGCGAGTGCAGTCCGTTGTCATCGAGCAGACATCTCT 166  
 791 MetAsnGluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerVal 810  
 165 ATGAATGAGAGCAGCAGCAGCTGTTTGAATCTTCTTCTGACATCTCTGCGTCCAGATGTC 106  
 811 VallylsileGlyAspArgCysThrGlnCysGlnGlyLeuProGlnGlySerSerLeu 830  
 105 GTAAGAATGTGTGACAGTGCTATACGACATGCCAGGCGATCCCCCAGGCGCTCCAGCCTA 46  
 831 SerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLys 845  
 45 TCCACCTGCTCTGACATCTGTGTTTGGAGACATGGAGACAAAG 1

BM824748 492 bp mRNA linear EST 06-MAR-2002  
 US K-EST0096335 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-99-E07  
 5', mRNA sequence.  
 ESSION BM824748  
 SION BM824748.1 GI:19181161  
 WORDS EST.  
 RGNISM Homo sapiens (human)  
 RCE Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 492)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 UNPUBLISHED  
 CONTACT: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 99 row: B column: 07  
 High quality sequence stop: 492.  
 Location/Qualifiers  
 1. 492  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S22SNU16n1-99-307"  
 /sex="F"  
 /tissue\_type="Ascites"  
 /cell\_type="lymphoblast-like"  
 /cell\_line="SNU-16"  
 /lab\_host="DH10B"  
 /clone\_lib="S22SNU16n1"  
 /note="Organ: Stomach; Vector: pT7T3-Pac; Site: 1. EcoRI;  
 Site 2: NotI; The S22SNU16 library was contributed by the  
 Soares laboratory and it was constructed as described by  
 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
 Research 6(9): 791-806. RNA was prepared from harvested  
 cells of SNU-16 culture. SNU-16 cell was obtained from  
 Korean Cell Line Bank (KCLB). SNU-16 was established from  
 ascitic fluids of Korean patients by Park J.G. et al.  
 (1990), Cancer Res 50: 2773-2780."

UT 10  
 24748  
 US  
 INITIATION K-EST0096335 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-99-E07  
 5', mRNA sequence.  
 ESSION BM824748  
 SION BM824748.1 GI:19181161  
 WORDS EST.  
 RGNISM Homo sapiens (human)  
 RCE Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 492)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 UNPUBLISHED  
 CONTACT: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 99 row: B column: 07  
 High quality sequence stop: 492.  
 Location/Qualifiers  
 1. 492  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S22SNU16n1-99-307"  
 /sex="F"  
 /tissue\_type="Ascites"  
 /cell\_type="lymphoblast-like"  
 /cell\_line="SNU-16"  
 /lab\_host="DH10B"  
 /clone\_lib="S22SNU16n1"  
 /note="Organ: Stomach; Vector: pT7T3-Pac; Site: 1. EcoRI;  
 Site 2: NotI; The S22SNU16 library was contributed by the  
 Soares laboratory and it was constructed as described by  
 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
 Research 6(9): 791-806. RNA was prepared from harvested  
 cells of SNU-16 culture. SNU-16 cell was obtained from  
 Korean Cell Line Bank (KCLB). SNU-16 was established from  
 ascitic fluids of Korean patients by Park J.G. et al.  
 (1990), Cancer Res 50: 2773-2780."

96 a 152 c 131 g 113 t

gment Scores:  
 id. No.: 2.77e-39 Length: 492  
 re: 558.00 Matches: 109  
 cent Similarity: 80.27% Conservative: 9  
 it Local Similarity: 74.15% Mismatches: 29

Query Match: 9.46% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-042-460-2 (1-1122) x BM824748 (1-492)  
 QY 898 ValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaProTyGlnLeuProAla 917  
 DB 1 GTGAACCTCCCTGTAGAACAGCAGGCGCTGGGTGGCGCTTTTGTTCAGATGCCGCGC 60  
 QY 918 HisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThrLeuGluValPheCys 937  
 DB 61 CACGGCGCTATTCCCTGGTGGCGCGCTGCTGCTGGATACCCGACCCCTGGAGGTGCAGAGC 120  
 QY 938 AspTyrSerGlyTyrAlaGlnThrSerIleTyrThrSerLeuThrPheGlnSerValPhe 957  
 DB 121 GACTACTCCAGCTATGCCCGGACCTCCATCAGACGACGCTCACCTTCAACCGCGGCTTC 180  
 QY 958 LysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArgLeuLysCysHisGly 977  
 DB 181 AAGCGCTGGGAGGAACATGCGTGCCTCAACTCTTTGGGTGCTTGGCGCTGAAGTGTCCAGC 240  
 QY 978 LeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLysIle 997  
 DB 241 CTGTTTCTGGATTTCAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTACAGATC 300  
 QY 998 PheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArg 1017  
 DB 301 CTCCTGCTGCAGCGGTACAGGTTTCAGCATGTGTGTGCAGCTCCATTTTCATCAGCAA 360  
 QY 1018 ValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysCysTyr 1037  
 DB 361 GTTTGGAAGAACCCCACTTTTCTGCGGCTCATCTCTGACACGGCGCTTCTGCTTGTCTAC 420  
 QY 1038 AlaIleLeuLysValLysAsn 1044  
 DB 421 TCCATCTCTGAAGCCCAAGAAC 441

RESULT 11  
 BU452535  
 LOCUS 603767927F1 CSEQREN14 Gallus gallus cDNA clone CHEST696f7 5', mRNA  
 DEFINITION sequence.  
 ACCSSION BU452535  
 VERSION BU452535.1 GI:25941846  
 KEYWORDS SSR.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 753)  
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.U.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 MEDLINE  
 PUBMED 12445392  
 COMMENT  
 Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST  
 )  
 PO Box 98, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. 753  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer"  
 /db\_xref="taxon:9031"  
 /clone="CHEST696f7"  
 /sex="Female"

FEATURES  
 source



```

/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="GSEQRBNI4"
/notes="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this in first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) 91: 9228-9232 and Bonaldo et al., Genome Research 6
reannealing hybridization was used."
E COUNT      230 a 155 c 132 g 236 t
GIN

gment Scores:
re: 1.98e-37 Length: 753
540.50 Matches: 106
cent Similarity: 76.36% Conservative: 20
t Local Similarity: 64.24% Mismatches: 36
ry Match: 9.16% Indels: 3
13 Gaps: 1

09-042-460-2 (1-1122) x BU452535 (1-753)
961 LysThrMetArgAnLysLeuLeuSerValLeuArgLeuLysCysHisGlyLeuPheLeu 980
5 AAAAAACATGAATGCAATGACGTGAGTCTCTCAACATGAATGCCATCCTTTACTTCT 64
981 AspLeuGlnValAnSerLeuGlnThrValCysIleAsnIleThrLysLeuPheLeuLeu 1000
65 GACTTAAAGATCAACAGCTTCAGACAGTCTCTTAAATCAATCAACAGATATTTTACTT 124
1001 GlnAlaThrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArgValArgLys 1020
125 CAGGCTTACAGTTCCATGCGTGTGTTCTTCTAGGTAGCAAGATCCTCTGATGTTCCCTTT 304
1021 AsnLeuThrPheLeuGlyIleIleSerSerGlnAlaSerCysCysThrAlaIleLeu 1040
185 AATCTGTATTTCTTCTAAGATCATCTCTGATCATCTCTCATGCTGCTATTTTATCTCG 244
1041 LysValLysAsnProGlyMetThrLeu-----LysAlaSerGlySerPheProPro 1057
245 AAAGCTAAATCCAGAGTTCTTCTAGGTAGCAAGATCCTCTGATGTTCCCTTT 304
1058 GluAlaAlaHisTrpLeuCysTyThrGlnAlaPheLeuLysLeuAlaHisSerVal 1077
305 GAGGCAGCAGAAATGCTGTGTACCATGCTCTCATGTCATCAACTGTCACCAACCAAGTT 364
1078 IleTVLysCysLeuLeuGlyProLeuArgThrAlaGlnLysLeuLeuCysArgLysLeu 1097
365 ATTACAAATGCTTACTTAAGCCCTTAAAGTCTTAAGATGATGATGTTGGGAGATC 424
1098 ProGluAlaThrMetThrIleLeuLeuAlaAlaAlaAspProAlaLeuSerThrAspPhe 1117
425 CCAGGGATATATGGAATGCTGAGCAGCGTGACGGAACCATCGTGTGTCAGATTTTC 484
1118 GlnThrIleLeuAsp 1122
485 AAAACTACTGGAC 499

```

```

MULT 12
49368
US
INITIATION
US musculus cDNA clone L93026124 5', mRNA sequence.
SESSION
USION
BY149368.1 GI:26285897

```

# KEYWORDS SOURCE ORGANISM

## REFERENCE AUTHORS

```

Mus musculus (house mouse)
Mus musculus
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 326)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quachobush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K.M., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons
P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perle, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring
B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, W.S., Teasdale
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander
E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12468851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami
M., Waki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1..326
/organism="Mus musculus"

```

## FEATURES source

```

/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="1930261124"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"
E COUNT      58 a  92 c  98 g  78 t
GIN

gment Scores:
d. No.:      2.15e-35      Length:      326
re:          511.00      Matches:      103
cent Similarity: 99.04%      Conservative: 0
t Local Similarity: 99.04%      Mismatches: 0
ry Match:    8.66%      Indels:      1
          13      Gaps:      0

```

09-042-460-2 (1-1122) x BY149368 (1-326)

```

67 AlaaspLeuSerPheHisGlnValSerSerLeuLeuValAlaArgValValGln 86
2 GCCACCTTTCTTCCACCGGTGTCATCCCTGAAAGAGCTGGTGCACGGTTGTGCAG 61
87 ArgLeuCysGluArgAenGluArgAenValLeuAlaPheGlyPheGluLeuLeuAsnGlu 106
62 AGATCTTCGCGCGCACGAGAGAAAGCTGCTGGCTTTGGCTTTGAGCTGCTTAACGAG 121
107 AlaArgGlyGlyProMetAlaPheThrSerSerValArgSerTyrLeu-ProAsnTh 126
122 GCCAGAGCGCGGCTCCCACTGGGCTTCACTAGTAGGCTGCTAGTACTTGCCTCCCAACAC 181
126 rValLeuLeuThrLeuArgValSerGlyAlaTrpMetLeuLeuLeuSerArgValGlyAs 146
182 TGTTATTGAGACCTCGCTGCTAGTGTGATGATGCTACTGTTGAGCGGAGTGGCGA 241
146 pAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeuLeuValProSerCy 166
242 CGACTGCTGCTACCTGCTGCGCACACTGCTCTTTATCTTCGTGCCCGCCAGCTG 301
166 SalATyGln 169
|||||
302 TGCCTACCAG 311

```

ULT 13  
88013

```

US      BI388013      775 bp      mRNA      linear      EST 20-MAY-2003
INITIATION BFL26_002716 Amphioxus 26hr cdna library (Name convention: BFL26 or
MEMGp531) Branchiostoma floridae cdna clone MEMGp531A0937 5', mRNA
sequence.
ESSION    BI388013      1 GI:30922834
SION      BI388013.1
WORDS     EST.
RCE       Branchiostoma floridae (Florida lancelet)
ORGANISM  Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
BRENCE    1 (bases 1 to 775)
AUTHORS   Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A., Herwig
,R., Vingron,M. and Lehrach,H.
New evidence for genome wide duplications at the origin of
vertebrates using an amphioxus gene set and completed animal
genomes
Genome Res. 13 (6), (2003) In press
JOURNAL   Contact: Panopoulou G
MENT       Laboratory 145, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhest-63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: panopoul@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting

```

(ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/tfide> )amphioxus/main.html Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTTACACTTTATGTTCCGGCTCG 3' (M13BSP)

BACKWARD: 5' GCTATTACGCGAGCTGGCGAAAGGGGATGTG 3' (M13PSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGTCGCGAATTCGCCGGT-3' pSPORT3/86

High quality sequence stop: 775.

#### FEATURES

```

Location/Qualifiers
1..775
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MEMGp531A0937"
/tissue_type="whole embryo"
/dev_stage="26 hrs (neurula stage)"
/lab_host="E.coli, XL1 blue"
/clone_lib="Amphioxus 26hr cdna library (Name convention:
BFL26 or MEMGp531)"
/note="Vector: pSPORT1 (Gibco BRL); Site 1: SalI, KpnI,
EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT
primed and directionally cloned in pSPORT1 vector using a
NotI (5'-pGACTAGTTCTAGATCGCGAGCGGCGGCC (T)15-3' and a
SalI 5'-TCGACCCACGCTCGG-3' adapters (Gibco BRL)."
BASE COUNT 176 a 217 c 194 g 186 t 2 others
ORIGIN
Alignment Scores:
Pred. No.:      3.14e-33      Length:      775
Score:          493.50      Matches:      110
Percent Similarity: 56.47%      Conservative: 34
Best Local Similarity: 43.14%      Mismatches: 104
Query Match:    8.36%      Indels:      7
DB:             12      Gaps:      3
US-09-042-460-2 (1-1122) x BI388013 (1-775)

```

```

Qy      810 valVallylleGlyAspArgCysTyrThrGlnCysGlnGlylleProGlnGlySerSer 829
Db      24 GTGATCCAGATCGGGGGAGCTACTACCGCGAGATCAGCGGGATCAGTCAGGGATCTCTGC 83
Qy      830 LeuSerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGlu 849
Db      84 CTGTCACACCTGCTCTAGTCTTCTTACGCTAATGAGAGAGACATCTCTGTGGC 143
Qy      850 valGlnArgAspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHis 869
Db      144 ATGACATCATGATGGGTTGCTGCTGCTGTGGTATGACTTCTCTGTTGACCCACAC 203
Qy      870 LeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCys 889
Db      204 CTGGACACGCTACAGCCTTCTCTGTACCATGTTGGATGATGATCCCTGAATATGGTGC 263
Qy      890 MetIleAsnLeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGly 909
Db      264 AGCGTTTCATCCGACAAAGTCTGTACCAACTTCCCTGCTCCAGCAC-AGGAAT 314
Qy      910 AlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuAsp 929
Db      315 GCTGTGGTGACATGTCAGCAGCATCGGCTGGTGTTCCTGGTGGCATCTCTTTCAC 374
Qy      930 ThrGlnThrLeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThr 949

```

[illegible]

**NOTES**  
 source  
 1. 880  
 Location/Qualifiers  
 /organism="Gallus gallus"  
 /mol\_type="rRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="CBST7987"  
 /sex="Female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCNN74"  
 /note="Organ: kidney + adrenal; Vector: pBluescript II  
 KS(+); Site 1: EcoRI; Site 2: NotI; This normalized  
 library was constructed from 1 million independent clones.  
 cDNA synthesis was initiated using an oligo(dT) primer.  
 Following this first strand reaction, double-stranded cDNA  
 was blunt-ended, ligated to NotI adapters, digested with EcoRI  
 , size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT	276 a	187 c	162 g	254 t	1 others
ORIGIN					

Alignment Scores:

***GameTree loaded.	
Pred. No.:	3,318-32
Score:	483.00
Length:	880
Matches:	110
Conservative:	24
Percent Similarity:	73.2%
Mismatches:	42
Best Local Similarity:	60.1%
Query Match:	8.19%
Indels:	8
DB:	13
DB:	2
DB:	2

US-09-042-460-2 (1-1122) x BU377259 (1-880)

942 TyrA]aG]nThrSerIleIysThrSerLeuThrPheGlnSerValpHeIysAlaGlyLys 961

260 TATTCCTTTTCTTCTATTCAAGTACAACTGCTGGGAAA 327  
 :  
 :

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

[illegible][illegible]

Qy 982 LeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLysIlePheLeuLeuGln 1001

D<sub>b</sub> 388 TTAAAGATCAACAGAGCCCTTCAGACAGTTCTAATTAAACATCTACAAGATATTTTACTTCAG 447

Qy 1002 AlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArgValArgLysAsn 1021

D<sub>b</sub> 448 GCTTACAGGTTCCATGCCCTGTGTTCTTCAGCTTCCATTCAACCAGAAAGTTAGGAATAAT 507

1022 LeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysCysTyrAlaIleLeuLys 1041

508 CCTCA TTTCTTCTCTCTTACCATCA TCTCTGAT TCTGCTTCA TGTCTGCTA TTTTAT TCTCTGAA 567

1058

[illegible][illegible]

```
QY      |  
I059 AlaaAaH8TpLEUCySiYrGlnAtApHeLeUeUtySleUAIAaRa-His-SerValL 107
```

Db 627 GCAGCAGAAATGGCTGTGCTACCATGCCTTCATTGTCAAACTGTCCAAACCAAAAGTAA 686

QY 1078 leTyrLysCysLeuLeuGlyProLeuArgThrAla-GlnLysLeuLeuCysArgLysLeu 1097

Db  
687 TTTACAAATGCTTACTTAAGCCCTTAAGCTCTATAGATGCATCTGTTGGCG--AAGATC 743

1098 ProGluAlaThrMetThrIleLeuLysAlaAlaAspProAlaLeuSer<sup>Thr</sup>AspPhe 1117

744 CCACGGATACCTTATCGGRACTGGTGAAGACGGGACCGAACCAATCGCTTGGTCCAGAAATTC 803

6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529

3  
3  
4  
4  
5  
5  
5

RESULT 13  
AA281296

LOCUS	AA281296	389 bp	trna	linear	EST I4-AUG-1997
DEFINITION	z408C02 r1-NCI	CGAP:GCB1	Homo sapiens	cDNA clone	IMAGE:7125625

mRNA sequence.

VERSION AA281296.1 GI:1924194

KEYWORDS  
EST.  
Homo sapiens (human)  
SOURCE

ORGANISM      Homo sapiens  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 2187 Std Error: 0.00  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 385.  
 Location/Qualifiers

## FEATURES

source  
 1..389  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:712562"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP GCB1"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD20+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 [5'-TGTTACATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'  
 ]. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

E COUNT 87 a 102 C 123 g 77 t  
 GIN

## Segment Scores:

d. NO.: 5.3e-31 Length: 389  
 re: 463.00 Matches: 87  
 cent Similarity: 82.03% Conservative: 18  
 t Local Similarity: 67.97% Mismatches: 23  
 ry Match: 7.85% Indels: 0  
 Gaps: 0

09-042-460-2 (1-1122) x AA281296 (1-389)

532 AlatrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePhe 551  
 1 GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGTCGAGCTGCTCAGGCTTTCTTT 60  
 552 TyrileThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrp 571  
 61 TATGTCCGGAGACCAAGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGG 120  
 572 SerLysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeu 591  
 121 AGCAAGTTGCCAAGCATTTGGAATCAGACAGACACTTGAAGAGGGTGCAGCTGCGGACGTG 180  
 592 SerGlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeu 611  
 181 TCGAAGCAGAGAGTCCAGCAGCATCCGGAAACCCAGCGCCCTGCTGACGTCACGACTC 240  
 612 ArgPheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGly 631  
 241 CGCTTCATCCCAAGCTTGACGGGCTCGGCCCGGATTTGTGAATGACTACGTCGTGGGA 300  
 632 ThrArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeu 651  
 301 GCCAGACGTTCCGACAGAGAAAGAGGCGCGAGCGCTCTACCTCGAGGGTGAAGGCACTG 360  
 652 PheSerMetLeuAsnTyrGluArg 659  
 361 TTCAGCGTGTCAACTACGACGG 384

Search completed: February 1, 2004, 19:29:33  
 Job time : 5272 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: February 2, 2004, 09:12:03 ; Search time 41 Seconds

(without alignments)  
933.003 Million cell updates/sec

le: US-09-780-669-114

uence: 1 MQCFSPFKTWMLFNLLIFL.....IGGLELAAMIVSMYLYCNLQ 241

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1258	100.0	241	19	AAW71870
2	1258	100.0	241	19	AAW59954
3	1258	100.0	241	19	AAW69386
4	1258	100.0	241	19	AAW58380
5	1258	100.0	241	21	AAW82003
6	1258	100.0	241	22	AAU69764
7	1258	100.0	241	22	AAU04962
8	1258	100.0	241	22	AAU01118
9	1258	100.0	241	22	AAG99003

10	1258	100.0	241	22	AAW74801	Prostate tumour an
11	1258	100.0	241	22	AAW90679	Human AK45-4 prot
12	1258	100.0	241	23	ABG94412	Human prostate tum
13	1258	100.0	241	23	ABG76666	Prostate tumour pr
14	1258	100.0	241	23	ABG95223	Human N1-1862 prot
15	1258	100.0	241	24	ABU71654	Prostate cancer as
16	1254	99.7	241	23	ABG96434	Human ovarian canc
17	1253	99.6	240	22	ABG63153	Human P503S invent
18	1241.5	98.7	240	23	ABW89113	Human polypeptide
19	1199	96.1	233	19	AAW75060	Human secreted pro
20	1189	95.3	273	19	AAW61618	Clone HPWAB25 of T
21	1182	94.0	258	23	ABG96433	Human ovarian canc
22	911	72.4	241	23	ABB09512	Human TSPAN-1-like
23	575	45.7	126	20	AAW35991	Extended human sec
24	526	41.8	121	23	ABP43214	Human ovarian anti
25	480	38.2	101	20	AAW13159	Human secreted pro
26	473	37.6	97	20	AAW12381	Human 5' EST seque
27	453	36.0	99	19	AAW47275	Human HPK-1A C4.8
28	452	35.9	89	20	AAW11862	Human 5' EST seque
29	407	32.4	103	23	ABP32558	Human glycoprotein
30	382	30.4	140	20	AAW86331	Kidney injury asso
31	373	29.7	108	20	AAW11883	Human 5' EST seque
32	368.5	29.3	248	22	AAW93282	Human protein sequ
33	368.5	29.3	248	22	AAW49503	Clone HCEIK90 #1.
34	368.5	29.3	248	23	ABP64749	Human protein SEQ
35	368.5	29.3	248	23	AAU79168	Human CD-81-like p
36	368.5	29.3	248	23	AAW48320	Human tetraspan.
37	361	28.7	249	23	AAU77412	Human NOV6 protein
38	351.5	27.9	236	22	AAW88457	Human membrane or
39	330.5	26.3	247	22	AAW47215	Human NOV7 protein
40	310	24.6	239	22	AAW93880	Human stomach canc
41	310	24.6	239	22	AAW93885	Human protein sequ
42	297	23.6	238	19	AAW61622	Clone HTPBA27 of T
43	296	23.5	164	22	AAW49509	Clone HCEIK90 #2.
44	288	22.9	96	22	AAW75719	Human colon cancer
45	288	22.9	252	19	AAW61624	Clone HPFEK40 of T

#### ALIGNMENTS

##### RESULT 1

AAW71870  
ID AAW71870 standard; Protein; 241 AA.  
XX AC AAW71870;  
XX DT 06-JAN-1999 (first entry)  
XX DE Amino acid encoded by prostate tumour clone N1-1862.  
XX KW Prostate; cancer; tumour; vaccine; immunogen; clone.  
XX OS Homo sapiens.  
XX PN W09837093-A2.  
XX PD 27-AUG-1998.  
XX PF 25-FEB-1998; 98WO-US03492.  
XX PR 09-FEB-1998; 98US-0020956.  
XX PR 25-FEB-1997; 97US-0806099.  
XX PR 01-AUG-1997; 97US-0904804.  
XX (CORI-) CORIXA CORP.  
XX Dillon DC, Xu J;  
XX DR WPI; 1998-609886/51.  
XX DR N-PEDB; AAW61202.  
XX PT Polypeptides comprising immunogenic portions of prostate proteins -

used in a vaccine for the treatment of prostate cancer

Example 1; Page 84-85; 130pp; English.

The present sequence is an immunogenic portion of a prostate tumour protein. The immunogen, or the DNA encoding it, can be used as a vaccine for the treatment of prostate cancer. The immunogen was isolated from a prostate tumour cDNA library obtained by subtracting a prostate tumour cDNA expression library with a normal tissue cDNA library.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFIAEVAANAVALVYTTMAEHFLTL 120  
 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFIAEVAANAVALVYTTMAEHFLTL 120  
 121 LVVPAIKKDYGSQEDFTQVWNTTMTKGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 121 LVVPAIKKDYGSQEDFTQVWNTTMTKGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 181 NTANETCTKQKAHQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQKAHQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

SULT 2  
 W59954

AAW59954 standard; Protein; 241 AA.

AAW59954;

02-DEC-1998 (first entry)

Amino acid sequence of the human tumour-associated antigen.

Human; tumour-associated antigen; PRAT; stimulation; cell proliferation; antagonist; cancer; genetic defect; sickle cell anaemia; agonist; antibody; hybridisation; probe.

Homo sapiens.

WO9838310-A1.

03-SEP-1998.

27-FEB-1998; 98WO-US03953.

28-FEB-1997; 97US-0808148.

(INCY-) INCYTE PHARM INC.

Goli SK, Hillman JL;

WPI; 1998-481208/41.

N-PSDB; AAV54014.

Human tumour-associated antigen PRAT - useful for stimulating cell proliferation and screening for antagonists useful to treat or prevent cell proliferation disorders e.g. cancers

PS Disclosure; Fig 1A-1C; 54pp; English.

XX This is the amino acid sequence of the human tumour-associated antigen (PRAT) used in the method of the invention for stimulating cell proliferation and screening for antagonists useful to treat or prevent cell proliferation disorders such as cancer, and genetic defect e.g. sickle cell anaemia. The polypeptides can be combined with a suitable carrier in pharmaceutical compositions, and also used to screen for antagonists, agonists, and to generate antibodies. PRAT agonists can be added to a cell to stimulate cell proliferation by increasing or prolonging the activity of PRAT as above. The antagonists can be combined with a suitable carrier in pharmaceutical compositions, which can be administered to subjects to treat or prevent disorders associated with cell proliferation, especially cancers. Antibodies specific for PRAT may be used directly as antagonists, or indirectly as a targeting or delivery mechanism to bring pharmaceutical agents to PRAT-expressing cells. They are also useful to diagnose conditions or diseases characterised by PRAT expression and to monitor therapeutic interventions. The polynucleotide encoding PRAT, or complementary sequences, can be used to produce hybridisation probes, useful to detect polynucleotides or fragments encoding PRAT, e.g. to diagnose diseases relating to polypeptide expression or monitor PRAT regulation during therapeutic intervention.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 Db 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 QY 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFIAEVAANAVALVYTTMAEHFLTL 120  
 Db 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFIAEVAANAVALVYTTMAEHFLTL 120  
 QY 121 LVVPAIKKDYGSQEDFTQVWNTTMTKGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 Db 121 LVVPAIKKDYGSQEDFTQVWNTTMTKGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 QY 181 NTANETCTKQKAHQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 Db 181 NTANETCTKQKAHQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 QY 241 Q 241  
 Db 241 Q 241

RESULT 3

AAW69386  
 ID AAW69386 standard; Protein; 241 AA.

XX AAW69386;

XX 25-MAR-2003 (updated)

DT 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone N1-1862 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection; therapy.

XX Homo sapiens.

XX WO9837418-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US03690.

XX



25-FEB-1997; 97US-0806596.  
01-AUG-1997; 97US-0904809.  
09-FEB-1998; 98US-0020747.

(CORI-) CORIXA CORP.

Dillon DC, Xu J;

WPI; 1998-480805/41.  
N-PSDB; AAV58587.

Novel human prostate specific tumour protein and fragments - useful  
for detecting and treating prostate cancers

Example 1; Page 89-90; 141pp; English.

This sequence is encoded by a human prostate tumour specific gene, and  
can be used in the method of the invention. The method is for detecting  
prostate cancer comprises contacting a biological sample with an agent  
able to bind an immunogenic portion of a prostate protein (such as  
this protein sequence). An antibody which binds to an immunogenic  
portion of the prostate protein, and the method can be used to detect,  
monitor progression of, or treat prostate cancers. The antibody may  
also be conjugated to a therapeutic agent for use in therapy of prostate  
cancers.

(Updated on 25-MAR-2003 to correct PR field.)

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSTKTMWILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60

1 MQCFSTKTMWILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60

61 IAGVWVVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAALVAVVYTTMAEHFTL 120

61 IAGVWVVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAALVAVVYTTMAEHFTL 120

121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

241 Q 241

241 Q 241

ULT 4

58380

AAW58380 standard; Protein; 241 AA.

AAW58380;

14-SEP-1998 (first entry)

Human secreted protein AR415\_4.

AR415\_4; secreted protein; protein factor; human.

Homo sapiens.

Key Location/Qualifiers

Peptide 14..26

/note= "putative leader/signal peptide or  
transmembrane domain"

Protein 27..241

/label= Mat\_protein

XX WO9817687-A2.

XX 30-APR-1998.

XX 24-OCT-1997; 97WO-US19590.

XX 24-OCT-1997; 97US-0740274.

XX 25-OCT-1996; 96US-0740274.

XX (GEMY ) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Racie LA, Spaulding V, Treacy M;

XX WPI; 1998-261426/23.

XX N-PSDB; AAV30916.

XX Nucleic acid encoding secreted protein from human cells - useful,

XX e.g. as immunomodulator, antitumour agent, promoters of tissue

XX growth, haemostatic and thrombolytic agents etc.

XX Claim 8; Page 67-68; 114pp; English.

XX This polypeptide, designated AT415\_4, is a novel human secreted

XX protein. Its amino acid sequence was deduced from a full-length

XX AT415\_4 cDNA clone (see AAV30916) isolated from a human adult retina

XX cDNA library. The predicted amino acid sequence shows homology to

XX human AAM35252 and CO-029 tumour associated antigens, and computer

XX predictions suggest a potential transmembrane domain centered

XX around amino acid 100 of the protein. 11 Novel human secreted

XX proteins (see AAW58580-90) are claimed. These can be expressed in

XX recombinant host cells for analysis, characterization, diagnostic

XX or therapeutic use. They can also be used as tissue or mol.wt.

XX markers, to generate antibodies, and in interaction trap assays.

XX They may have biological activities, e.g. cytokine, immunomodulator,

XX haematopoiesis regulating activity, tissue growth activity, activin

XX or inhibin activity, chemotactic or chemokinetic activity,

XX haemostatic and thrombolytic activity, receptor/ligand activity,

XX antiinflammatory, cadherin and tumour invasion suppressor activity,

XX and tumour inhibition activity. The proteins can be expressed in

XX vivo from DNA, introduced in gene therapy vectors.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MQCFSTKTMWILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60

Db 1 MQCFSTKTMWILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60

Cy 61 IAGVWVVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAALVAVVYTTMAEHFTL 120

Db 61 IAGVWVVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAALVAVVYTTMAEHFTL 120

Cy 121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

Db 121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

Cy 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

Db 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

Cy 241 Q 241

Db 241 Q 241

RESULT 5  
AAW82003

AAV82003 standard; Protein; 241 AA.

AAV82003;

13-JUN-2000 (first entry)

Human immunogenic prostate tumour protein sequence SEQ ID NO:114.

Human; prostate cancer; diagnosis; tumour; gene therapy; detection; immunogenic; cytostatic; vaccine.

Homo sapiens.

WO200004149-A2.

27-JAN-2000.

14-JUL-1999; 99WO-US15838.

14-JUL-1998; 98US-0115453.

14-JUL-1998; 98US-0116134.

23-SEP-1998; 98US-0159812.

23-SEP-1998; 98US-0159822.

15-JAN-1999; 98US-0232149.

15-JAN-1999; 98US-0232880.

09-APR-1999; 99US-0288946.

(CORI-) CORIXA CORP.

Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;

WPI; 2000-171268/15.

New polypeptide useful for treating and diagnosing prostate cancer comprises an immunogenic portion of prostate tumor protein -

Claim 3; Page 139-140; 263pp; English.

The present invention describes isolated polypeptides, comprising an immunogenic portion of a prostate tumour protein (PTP). The polypeptides and polynucleotides encoding them have cytostatic activity and can be used in vaccines and in gene therapy. The polypeptides and polynucleotides encoding them, antigen presenting cells which express the polypeptides, antibodies against the polypeptides and vaccines comprising them can be used for inhibiting the development of prostate cancer in a patient. The polypeptides can be used to generate antibodies or anti-idiotypic antibodies for passive immunotherapy. A portion of the polynucleotides encoding the polypeptides can be used as a probe or to modulate the expression of the polypeptides. AAA06241 to AAA06691 and AAV82000 to AAV82020 represent sequences used in the exemplification of the present invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 21; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

61 IAGVVVVFALGFLGCGVATKESKCALVTFPFFILLIFIAEVAARVAVVTTMAEHLTL 120

61 IAGVVVVFALGFLGCGVATKESKCALVTFPFFILLIFIAEVAARVAVVTTMAEHLTL 120

121 LVVPAIKDYGSGEDFTQVWNTMTKGLKCCGFTNYTDFEDSPYKENSAPPPFCNDNVT 180

121 LVVPAIKDYGSGEDFTQVWNTMTKGLKCCGFTNYTDFEDSPYKENSAPPPFCNDNVT 180

181 NTANETCTQKARDQVEGCFNQLLYDTRNAVTVGGVAAIGGLELAAMTVSYLYCNL 240

181 NTANETCTQKARDQVEGCFNQLLYDTRNAVTVGGVAAIGGLELAAMTVSYLYCNL 240

QY 241 Q 241

Db 241 Q 241

RESULT 6

AAU69764

ID AAU69764 standard; Protein; 241 AA.

XX AAU69764;

DT 30-JAN-2002 (first entry)

DE Human prostate cDNA encoded protein #4.

KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.

OS Homo sapiens.

PN WO200173032-A2.

PD 04-OCT-2001.

PF 27-MAR-2001; 2001WO-US09919.

PR 27-MAR-2000; 2000US-0536857.

PR 09-MAY-2000; 2000US-0568100.

PR 12-MAY-2000; 2000US-0570737.

PR 13-JUN-2000; 2000US-0593793.

PR 27-JUN-2000; 2000US-0605783.

PR 10-AUG-2000; 2000US-0636215.

PR 29-AUG-2000; 2000US-0651236.

PR 06-SEP-2000; 2000US-0657279.

PR 02-OCT-2000; 2000US-0679426.

PR 10-OCT-2000; 2000US-0685166.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2001-639232/73.

DR N-PSDB; AAS63558.

XX New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer -

PS Claim 2; Page 270-271; 579pp; English.

XX The invention relates to isolated prostate-specific

CC polynucleotides, polypeptides, fusion proteins of the polypeptides,

CC antibodies raised against the polypeptides (or antigenic epitopes

CC derived from them) and antigen-presenting cells expressing the

CC polypeptides. The antibodies are useful for detecting the presence of

CC cancer, especially prostate cancer. The polypeptides, polynucleotides and

CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development

CC of cancer especially prostate cancer. Compositions comprising the

CC polynucleotide and/or polypeptide are useful for stimulating an immune

CC response, and for treating cancer. The oligonucleotide is useful for

CC detecting cancer. The present sequence is a prostate specific

CC polypeptide of the invention.

XX Sequence 241 AA;

SQ

Query Match 100.0%; Score 1258; DB 22; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

1 MQCFSEIKTMMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60  
 61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120  
 61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120  
 121 LVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180  
 121 LVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180  
 181 NTANETCTKQKAHQDKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQKAHQDKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

ULT 7  
 04962

AAU04962 standard; Protein; 241 AA.

AAU04962;

24-OCT-2001 (first entry)

Human prostate tumour protein N1-1862.

Human; prostate tumour protein; prostate cancer.

Homo sapiens.

US6262245-B1.

17-JUL-2001.

25-FEB-1998; 98US-0030607.

25-FEB-1997; 97US-0806099.

01-AUG-1997; 97US-0904804.

09-FEB-1998; 98US-0020956.

(CORI-) CORIXA CORP.

Xu J, Dillon DC;

WPI; 2001-440862/47.

N-PSDB; AAS10109.

Novel polynucleotide encoding polypeptide comprising a portion of prostate tumour protein useful for inhibiting development of prostate cancer or for treating prostate cancer in a patient -

Example 1; Column 128-129; 105pp; English.

The sequence is a partial prostate tumour protein, encoded by a prostate tumour specific cDNA. The DNA is useful for inhibiting the development of prostate cancer or for treating prostate cancer in a patient.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSEIKTMMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60

1 MQCFSEIKTMMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60

61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120

61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120

Qy 121 LVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180  
 Db 121 LVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180  
 Qy 181 NTANETCTKQKAHQDKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 Db 181 NTANETCTKQKAHQDKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 Qy 241 Q 241  
 Db 241 Q 241

RESULT 8

AAU01118

ID AAU01118 standard; Protein; 241 AA.

XX AC AAU01118;

DT 04-OCT-2001 (first entry)

XX Human prostate-specific amino acid sequence N1-1862.

XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 cytostatic; gene therapy; metastasis.

XX OS Homo sapiens.

XX WO200151633-A2.

XX 19-JUL-2001.

XX 16-JAN-2001; 2001WO-US01574.

XX 14-JAN-2000; 2000US-0483672.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

PI Xalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;

PI Wang A, Meagher MJ;

XX WPI; 2001-425873/45.

XX New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PT for use in vaccines -

XX Claim 2; Page 268-269; 543pp; English.

XX The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAU01115 to  
 CC AAU01118 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQCFSEIKTMMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60

||||| 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPJSSSAMQFVNVGYFL 60  
61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
121 LVVPAIKKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
121 LVVPAIKKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
241 Q 241  
241 Q 241

ULT 9  
399003  
AAG99003 standard; Protein; 241 AA.

AAG99003;

25-SEP-2001 (first entry)

Human prostate-specific amino acid sequence N1-1862/P503S.

Human; prostate cancer; therapy; diagnosis; cat eye syndrome;  
chromosome 22q11.2; prostate-specific protein; chromosome 1;  
prostate specific antigen; PSA.

Homo sapiens.

WO200134802-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US30904.

12-NOV-1999; 99US-0439313.

18-NOV-1999; 99US-0443686.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;

WPI; 2001-308785/32.

Isolated polypeptide comprising at least an immunogenic portion of a  
prostate-specific protein, useful in the diagnosis and therapy of  
prostate cancer -

Claim 3; Page 168-169; 325pp; English.

The present invention describes an isolated polypeptide (P1) comprising  
at least an immunogenic portion of a prostate-specific protein, or its  
variant. Also described are polynucleotides (N1) encoding (P1). (P1) and  
(N1) have cytostatic activity and can be used in vaccine production.  
The polypeptides, nucleic acids and antibodies from the present  
invention are useful in the diagnosis and therapy of prostate cancer.  
Prostate specific genes P704P, P712P, P774P, P775P and P305D are located  
in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome  
region. Prostate specific antigen (PSA) P501S was located on  
chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent  
polynucleotide and polypeptide sequences used in the exemplification  
of the present invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132; Indels 0; Gaps 0;  
Matches 241; Conservative 0; Mismatches 0;  
QY 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPJSSSAMQFVNVGYFL 60  
DB 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPJSSSAMQFVNVGYFL 60  
QY 61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
DB 61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
QY 121 LVVPAIKKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
DB 121 LVVPAIKKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
QY 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
DB 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
QY 241 Q 241  
DB 241 Q 241

RESULT 10

AAAB74801

ID AAB74801 standard; Protein; 241 AA.

AC AAB74801;

XX 14-JUN-2001 (first entry)

Prostate tumour antigen predicted amino acid sequence for N1-1862.

Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;  
Prostate cancer; immunogenic; cytostatic; vaccine.

OS Homo sapiens.

PN WO200125272-A2.

PD 12-APR-2001.

PF 04-OCT-2000; 2000WO-US27464.

PR 04-OCT-1999; 99US-0157455.

PA (CORI-) CORIXA CORP.

PI Xu J, Skeiky YAW, Reed SG, Cheever MA;

XX WPI; 2001-245062/25.

DR N-PSDB; AAH02531.

PT Prostate specific protein and its encoding polynucleotide, useful for  
the treatment and diagnosis of prostate cancer -

PS Claim 3; Page 158-159; 276pp; English.

The present invention describes an isolated polypeptide (I) comprising  
at least an immunogenic portion of a prostate tumour antigen protein or  
its variant. (I) have cytostatic activity and can be used in vaccine  
production. (I), prostate tumour antigen polynucleotides, an antigen  
presenting cell (APC e.g. a dendritic cell) that expresses (I), and a  
pharmaceutical composition containing (I) are useful for inhibiting the  
development of cancer in a patient. Antibodies specific for prostate  
specific proteins and oligonucleotides that hybridise to a  
polynucleotide that encodes a prostate specific protein are useful  
for detecting the presence or absence of a cancer or monitoring the  
progression the progression of a cancer, especially prostate cancer.  
AAH02442 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences  
used in the exemplification of the present invention.

Sequence 241 AA;  
 Query Match 100.0%; Score 1258; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWSDIGASFLKIFGFLSSAMQFVNVGYFL 60  
 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWSDIGASFLKIFGFLSSAMQFVNVGYFL 60  
 61 IAAGVVVFALGFLGCGYKATESKCALVTFFILLIFIAEVAAVVALVYTTMAEHFTL 120  
 61 IAAGVVVFALGFLGCGYKATESKCALVTFFILLIFIAEVAAVVALVYTTMAEHFTL 120  
 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNV 180  
 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNV 180  
 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

UT 11  
 90679

AA890679 standard; Protein; 241 AA.

AA890679;

07-JUN-2001 (first entry)

Human AR415\_4 protein sequence SEQ ID 35.

Human; secreted protein; nutrient; cytokine modulator; proliferation; differentiation; immune system modulator; tissue growth; chemotactic; haemostatic; thrombolytic; anti-inflammatory; tumour inhibition; haematopoiesis.

Homo sapiens.

WO200119388-A1.

22-MAR-2001.

14-SEP-2000; 2000WO-US25135.

17-SEP-1999; 99US-0398829.

(GENY) GENETICS INST INC.

Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C, Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;

WPI: 2001-244801/25.

N-PSDB; AAF98395.

Isolated nucleic acids encoding polypeptides, useful for modulating e.g. cytokine and cell proliferation/differentiation activity, the immune system and hematopoiesis regulating activity -

Disclosure; Page 403-404; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted proteins AAB90667 - AAB90750. The cDNA clones are isolated from various tissue types, and may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate protein expression. The polypeptides and nucleic acids may be used as nutrients or to modulate cytokine and cell proliferation/differentiation activity and may also be

CC involved in modulation of the immune system. The cDNA sequences, proteins, their agonists and/or antagonists exhibit haematopoiesis regulating activity; tissue growth activity; activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic and thrombolytic activity; receptor/ligand activity; anti-inflammatory activity; haematopoiesis activity; cadherin/tumour suppressor activity; and/or tumour inhibition activity. Included in the invention are probes represented in AAF98490 - AAF98572 which are specific for the cDNA clones encoding the secreted proteins.

SQ Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWSDIGASFLKIFGFLSSAMQFVNVGYFL 60

DB 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWSDIGASFLKIFGFLSSAMQFVNVGYFL 60

QY 61 IAAGVVVFALGFLGCGYKATESKCALVTFFILLIFIAEVAAVVALVYTTMAEHFTL 120

DB 61 IAAGVVVFALGFLGCGYKATESKCALVTFFILLIFIAEVAAVVALVYTTMAEHFTL 120

QY 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNV 180

DB 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNV 180

QY 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

DB 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

QY 241 Q 241

DB 241 Q 241

RESULT 12

ABG94412

ID ABG94412 standard; Protein; 241 AA.

XX ABG94412;

XX 27-NOV-2002 (first entry)

DE Human prostate tumour protein partial sequence #4.

XX Human; immunogenic; prostate protein; prostate tumour protein;

KW prostate cancer; cytostatic; vaccine.

XX Homo sapiens.

XX US2002030372-A1.

XX 11-JUL-2002.

XX 14-JUL-1998; 98US-0115453.

XX 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

PR 03-FEB-1998; 98US-0020956.

PR 25-FEB-1998; 98US-0030607.

XX (XUJJ/) XU J.

PA (DILL/) DILLON D C.

XX Xu J, Dillon DC;

XX WPI; 2002-642373/69.

DR N-PSDB; ABS21255.

XX Novel polypeptides useful as vaccines for inhibiting prostate cancer development, comprise an immunogenic portion of prostate protein -

Example 1; Page 61-62; 101pp; English.

The present invention relates to a new polypeptide comprising an immunogenic portion of a prostate protein. The invention is useful for inhibiting the development of prostate cancer in a patient. The invention is also useful as markers for diagnosing prostate cancer and for monitoring diseases progression in patients. The present amino acid sequence represents a human prostate tumour protein.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 23; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFTAEVAAAVVALVTTWAEHFLTL 120  
 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFTAEVAAAVVALVTTWAEHFLTL 120  
 121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
 121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

SULT 13

376666  
 ABG76666 standard; Protein; 241 AA.

ABG76666;

05-NOV-2002 (first entry)

Prostate tumour protein #4.

Human; prostate tumour; immunotherapy; prostate cancer.

Homo sapiens.

US2002081580-A1.

27-JUN-2002.

25-FEB-1998; 98US-0030606.

25-FEB-1997; 97US-0806596.

01-AUG-1997; 97US-0904809.

09-FEB-1998; 98US-0020747.

(XUJG/) XU J.

(DILL/) DILLON D C.

Xu J, Dillon DC;

WPI; 2002-607662/65.

Detecting prostate cancer comprises contacting a sample with an agent capable of binding to a polypeptide with an immunogenic portion of a prostate protein, oligonucleotide primers or a probe specific for DNA encoding the polypeptide.

PS Example 1; Page 67-68; 111pp; English.

XX The invention relates to a method of detecting prostate cancer by contacting a biological sample from a patient with: (a) a binding agent that binds to a polypeptide having an immunogenic portion of a prostate protein or its variant; (b) 2 oligonucleotide primers, where 1 of the oligonucleotides is specific for a DNA encoding the polypeptide of (a); or (c) an oligonucleotide probe specific for a DNA molecule encoding the polypeptide of (a). The method and polypeptides are useful for diagnosing, treating, particularly by immunotherapy, monitoring the progression, and inhibiting the development of prostate cancer in a patient. The polypeptides may be used to generate antibodies useful for the diagnosis and monitoring of prostate cancer. ABG76663-ABG76669 CC represent human prostate tumour protein sequences of the invention.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 23; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 DB 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 QY 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFTAEVAAAVVALVTTWAEHFLTL 120  
 DB 61 IAAGVVVFALGFLGCGYKTESKCALVTFFILLIFTAEVAAAVVALVTTWAEHFLTL 120  
 QY 121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
 DB 121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
 QY 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
 DB 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
 QY 241 Q 241  
 DB 241 Q 241

RESULT 14

ABB95223

ID ABB95223 standard; Protein; 241 AA.

XX ABB95223;

XX 19-JUL-2002 (first entry)

XX Human N1-1862 protein SEQ ID NO 114.

DE Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant; gene therapy.

XX Homo sapiens.

XX US2002022248-A1.

XX 21-FEB-2002.

XX 12-JAN-2001; 2001US-0759143.

XX 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

PR 09-FEB-1998; 98US-0020956.

PR 25-FEB-1998; 98US-0030607.

PR 14-JUL-1998; 98US-0115453.

PR 23-SEP-1998; 98US-0159812.

PR 15-JAN-1999; 99US-0232149.

PR 09-APR-1999; 99US-028846.

PR 13-JUL-1999; 99US-0352616.

PR 12-NOV-1999; 99US-0439313.

18-NOV-1999; 99US-0443686.  
 14-JAN-2000; 2000US-0483672.  
 27-MAR-2000; 2000US-0536857.  
 09-MAY-2000; 2000US-0568100.  
 12-MAY-2000; 2000US-0570737.  
 13-JUN-2000; 2000US-0593793.  
 27-JUN-2000; 2000US-0605783.  
 10-AUG-2000; 2000US-0636215.  
 29-AUG-2000; 2000US-0651236.  
 06-SEP-2000; 2000US-0657279.  
 02-OCT-2000; 2000US-0679426.  
 10-OCT-2000; 2000US-0685166.

(XUJ//) XU J.  
 (DILL//) DILLON D C.  
 (MITC//) MITCHAM J L.  
 (HARL//) HARLOCKER S L.  
 (JIAN//) JIANG Y.  
 (KALO//) KALOS M D.  
 (FANG//) FANGER G R.  
 (RETT//) RETTER M W.  
 (STOL//) STOLK J A.  
 (DAYC//) DAY C H.  
 (VEDV//) VEDVICK T S.  
 (CART//) CARTER D.  
 (LISX//) LI S X.  
 (WANG//) WANG A.  
 (SKEI//) SKEIKY Y A W.  
 (HEPL//) HEPLER W T.  
 (HEND//) HENDERSON R A.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

WPI; 2002-255649/30.

New prostate-specific polynucleotides for diagnosing and treating  
 diseases, in particular prostate cancer, and as markers for the  
 progression of cancer

Claim 2; SEQ ID NO 114; 87pp; English.

The present invention provides prostate-specific coding sequences and  
 their encoded proteins. These can be used in the diagnosis and treatment  
 of cancers, particularly prostate cancer. The present sequence is a  
 protein described in the invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 23; Length 241;  
 est Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MOCSEFTKMWLENAFLCGGALLAVGTVWSIDGASFLKIFGPISSAMOFVNVGYFL 60  
 1 MOCSEFTKMWLENAFLCGGALLAVGTVWSIDGASFLKIFGPISSAMOFVNVGYFL 60  
 61 IAAGVWVFALGFLGCGAKTESKCALVTPFFILLIFIAEVAADVVALVYTTMAEHFLTL 120  
 61 IAAGVWVFALGFLGCGAKTESKCALVTPFFILLIFIAEVAADVVALVYTTMAEHFLTL 120  
 121 LWVPAIKDYGSQEDFTQVWNTWMKGLKCGFNVTDFEDSPYFKENSAPFPFCNDNVT 180  
 121 LWVPAIKDYGSQEDFTQVWNTWMKGLKCGFNVTDFEDSPYFKENSAPFPFCNDNVT 180  
 181 NTANETCTKQAHQKQEGGCFNQLLYDIRNAVTVGVGAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQAHQKQEGGCFNQLLYDIRNAVTVGVGAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

RESULT 15  
 ABU71654

ID ABU71654 standard; Protein; 241 AA.

XX AC ABU71654;

XX DT 10-JUN-2003 (first entry)

XX DE Prostate cancer associated protein #3.

XX KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 immunogen; cancer; prostate specific antigen; PSA; membrane antigen;  
 PSMA.

XX OS Homo sapiens.

XX EN US2002192763-A1.

XX PD 19-DEC-2002.

XX PF 29-JUN-2001; 2001US-0895793.

XX PR 17-APR-2000; 2000US-157455P.

XX PR 04-OCT-2000; 2000US-0679272.

XX PR 28-MAR-2001; 2001US-0822827.

XX PA (XUJ//) XU J.

XX PA (DILL//) DILLON D C.

XX PA (MITC//) MITCHAM J L.

XX PA (HARL//) HARLOCKER S L.

XX PA (JIAN//) JIANG Y.

XX PA (KALO//) KALOS M D.

XX PA (FANG//) FANGER G R.

XX PA (RETT//) RETTER M W.

XX PA (STOL//) STOLK J A.

XX PA (DAYC//) DAY C H.

XX PA (VEDV//) VEDVICK T S.

XX PA (CART//) CARTER D.

XX PA (LISX//) LI S X.

XX PA (WANG//) WANG A.

XX PA (SKEI//) SKEIKY Y A W.

XX PA (HEPL//) HEPLER W T.

XX PA (HEND//) HENDERSON R A.

XX PA (HURA//) HURAL J.

XX PA (MCNE//) MCNEILL P D.

XX PA (HOUG//) HOUGHTON R L.

XX PA (DBAS//) Y DE BASSOLS C V.

XX PA (FOYT//) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 McNeill PD, Houghton RL, Y De Bassols CV, Foy TM;  
 WPI; 2003-352711/33.

XX New fusion protein comprising prostate-specific polypeptides, or its  
 immunogenic portions, useful for diagnosing, preventing and/or treating  
 cancer, particularly prostate cancer

XX Example 1; SEQ ID NO 114; 85pp; English.

The invention describes a fusion protein comprising at least one amino  
 acid sequence of immunogenic portions of any of the 3 sequences not  
 defined in the specification, or sequences having at least 70 or 90 %  
 sequence identity to any one of the 35 sequences defined in the USPTO  
 web site, which is encoded by any of the 4 nucleotide sequences not  
 defined in the specification. The fusion protein, composition and  
 methods are useful for diagnosing, preventing and/or treating cancer,  
 particularly prostate cancer. The proteins are useful as markers to



indicate the presence or absence of cancer. This is the amino acid sequence of a prostate cancer therapy associated protein.  
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at  
seqdata.uspto.gov/sequence.html?DocID=US20020192763.

```
Sequence 241 AA;
Query Match 100.0%; Score 1258; DB 24; Length 241;
Est Local Similarity 100.0%; Pred. No. 6.4e-132;
atches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MQCPSEFKTWMLFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLSSAMQFVNVGYFL 60
1 MQCPSEFKTWMLFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLSSAMQFVNVGYFL 60
61 IAAGVWVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAAVVALVYTTMAEHFLTL 120
61 IAAGVWVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAAVVALVYTTMAEHFLTL 120
121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180
121 LVVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDNVT 180
181 NTANETCTKOKAHDQKVEGCFNOLLYDIRTNVTVGVGAAGIGGLELAAMIVSMYLYCNL 240
181 NTANETCTKOKAHDQKVEGCFNOLLYDIRTNVTVGVGAAGIGGLELAAMIVSMYLYCNL 240
241 Q 241
241 Q 241
```

rch completed: February 2, 2004, 09:18:03  
Time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 4, 2004, 14:09:14 ; Search time 47 Seconds  
(without alignments)  
3789.175 Million cell updates/sec

le: US-09-042-460-2  
fect score: 5901  
uence: 1 NTRAPRCPAVALSLRSRYRE.....TILKARADPALSTDFQTILD 1122

ring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

urched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

imum DB seq length: 0  
imum DB seq length: 2000000000

it-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : A\_Geneseq 19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

alt No.	Score	Match	Length	DB ID	Description
1	5901	100.0	1122	20	AA1980
2	5854	99.2	1122	20	AA1981
3	3505	59.4	1132	19	AAW46957
4	3505	59.4	1132	20	AA1982
5	3505	59.4	1132	20	AA1983
6	3505	59.4	1132	20	AA1984
7	3505	59.4	1132	20	AA1985
8	3505	59.4	1132	20	AA1986
9	3505	59.4	1132	22	AA1987

10	3505	59.4	1132	22	AA1988
11	3505	59.4	1132	22	AA1989
12	3505	59.4	1132	22	AA1990
13	3505	59.4	1132	23	AA1991
14	3505	59.4	1132	23	AA1992
15	3505	59.4	1132	24	AA1993
16	3505	59.4	1154	19	AAW61350
17	3505	59.4	1154	19	AAW47008
18	3499	59.3	1285	19	AAW47000
19	3498	59.3	1132	19	AAW71376
20	3498	59.3	1132	20	AA1994
21	3498	59.3	1132	20	AA1995
22	3498	59.3	1132	20	AA1996
23	3498	59.3	1132	21	AA1997
24	3496	59.2	1132	19	AAW56113
25	3486.5	59.1	1199	19	AAW47007
26	3471	58.8	1166	20	AAW00647
27	3463	58.7	1405	19	AAW56101
28	3435	58.2	1120	20	AA1998
29	3426	58.1	1120	20	AA1999
30	3320	56.3	1150	19	AAW47006
31	3277.5	55.5	1053	20	AA1990
32	3239.5	54.9	1093	20	AA1991
33	3198.5	54.2	1041	20	AA1992
34	3198.5	54.2	1041	20	AA1993
35	2881.5	48.8	948	20	AA1994
36	2877.5	48.8	948	20	AA1995
37	2814.5	47.7	936	20	AA1996
38	2805.5	47.5	936	20	AA1997
39	2790.5	47.3	949	19	AAW61349
40	2195	37.2	807	19	AAW46997
41	2195	37.2	807	20	AA1998
42	2195	37.2	807	20	AA1999
43	2132	36.1	622	20	AA1990
44	2073	35.1	591	20	AAW57384
45	1709	29.0	499	24	AA1991

## ALIGNMENTS

RESULT 1  
AA1980  
ID AA1980 standard; Protein; 1122 AA.  
XX  
AC AA1980  
DT 13-SEP-1999 (first entry)  
XX  
DE Murine telomerase reverse transcriptase (mTERT) enzyme.  
XX  
XX Telomerase reverse transcriptase; TERT; mouse; telomere length assay;  
KW immunogen; enzyme; telomerase-mediated DNA replication.  
XX  
XX Mus sp.  
XX  
XX WO9927113-A1.  
XX  
XX 03-JUN-1999.  
XX  
XX 25-NOV-1998; 98WO-US25211.  
XX  
XX 16-MAR-1998; 98US-0042460.  
XX  
XX 26-NOV-1997; 97US-0979742.  
XX  
XX (GERO-) GERON CORP.  
XX (YESH) UNIV YESHIVA EINSTEIN COLLEGE.  
XX  
XX Allisopp R, Depinho R, Greenberg R, Morin GB;  
XX WPI, 1999-347722/29.  
XX N-PSDB; AA1980994.

Human telomerase p  
Human protein #2.  
Heart muscle cell  
Human telomerase r  
Human telomerase r  
Human telomerase r  
Human telomerase p  
Human telomerase p  
Glutathione-S-tran  
His tagged thior  
Human telomerase c  
Human ESR2 protein  
Human telomerase p  
Truncated telomera  
hEST2, a human tel  
Human telomerase r  
Glutathione-S-tran  
Telomerase (ver. 2  
Telomerase protein  
Telomerase (ver. 2  
Glutathione-S-tran  
Altered C-terminus  
Altered C-terminus  
Altered C-terminus  
N-terminal truncat  
Truncated telomera  
Truncated telomera  
Human telomerase p  
Human telomerase r  
N-terminal truncat  
Truncated telomera  
Human CRT-1 protei  
A catalytic telome  
Amino acid sequenc

Mouse telomerase reverse transcriptase (mTERT) enzyme proteins and nucleic acids

Claim 8; Fig 2; 135pp; English.

The invention relates to a mouse telomerase reverse transcriptase (mTERT) enzyme. Compositions containing mTERT can be used in telomere length assays. Isolated mTERT is useful as an immunogen for the production of monoclonal or polyclonal antibodies. The method is useful for assessing the degree of purification and identification of new mTERT species, such as an mTERT allele, homolog or isoform, or to screen for modulators (antagonists and agonists) of telomerase-mediated DNA replication. Other telomerase enzymes such as human TERT (hTERT). The present sequence represents a mTERT enzyme.

Sequence 1122 AA;  
Query Match 100.0%; Score 5901; DB 20; Length 1122;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MTRAPCPAVRSLLRSRYREVWPLATFVRRLQPEGRRLVQGPDKIYRTLVAQCLVCMHW 60  
1 MTRAPCPAVRSLLRSRYREVWPLATFVRRLQPEGRRLVQGPDKIYRTLVAQCLVCMHW 60  
61 GSOPPPADLSFHQVSSIKELVAVQVRLCERNVNLAFGEELLNARGGPPMAFTSSVR 120  
61 GSOPPPADLSFHQVSSIKELVAVQVRLCERNVNLAFGEELLNARGGPPMAFTSSVR 120  
121 SYLPNTVIETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPNTVIETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASYPTRVGNFNTLRFLOQIKSSROBAPKPLALPSRGTKEHLSLTSTS 240  
181 TTDIWPVSASYPTRVGNFNTLRFLOQIKSSROBAPKPLALPSRGTKEHLSLTSTS 240  
241 VPSAKKARCPVPRVEGPHRQVLPFPSCGSKSWPSPARPEVPTAEKDISKGVSDLSL 300  
241 VPSAKKARCPVPRVEGPHRQVLPFPSCGSKSWPSPARPEVPTAEKDISKGVSDLSL 300  
301 SGVCCCKHPSSTLSLPPQNAFARLPFIETHFLYSRGDQERLNPSFLLSNLQPNLT 360  
301 SGVCCCKHPSSTLSLPPQNAFARLPFIETHFLYSRGDQERLNPSFLLSNLQPNLT 360  
361 GARLVEIIFLGSRPRTSGPLCTRHSRYWOMRPLFOQLLVNHAECQVRLLSHCRF 420  
361 GARLVEIIFLGSRPRTSGPLCTRHSRYWOMRPLFOQLLVNHAECQVRLLSHCRF 420  
421 RTAQOQVTDALNTSPPLMDLLRLHSSPWQVYGLRACLCVKVVSASLWTRNRRFFN 480  
421 RTAQOQVTDALNTSPPLMDLLRLHSSPWQVYGLRACLCVKVVSASLWTRNRRFFN 480  
481 LKRTISLCKYKLSLQELMKMKVCHWLRSSPGKDRVPAEHLRLERILATFLWMD 540  
481 LKRTISLCKYKLSLQELMKMKVCHWLRSSPGKDRVPAEHLRLERILATFLWMD 540  
541 TVVQQLRSFFYITESTFQKNLFFYKSVWSLQSIGVRQHLERVRLRELSQEBVRHQ 600  
541 TVVQQLRSFFYITESTFQKNLFFYKSVWSLQSIGVRQHLERVRLRELSQEBVRHQ 600  
601 DTWLAMPICRLRFTPKNGRLPIVNSYMGTRALGRKQAQHTORLXLTLFSLMNYERT 660  
601 DTWLAMPICRLRFTPKNGRLPIVNSYMGTRALGRKQAQHTORLXLTLFSLMNYERT 660  
661 KPHLMGSSVLGMDIIVTRWAFVLRVLRALDQTPRMVFKADVTGAYDAIPQGLVEVVA 720  
661 KPHLMGSSVLGMDIIVTRWAFVLRVLRALDQTPRMVFKADVTGAYDAIPQGLVEVVA 720  
721 NMRHSESTYCIRQYAVVRDSDQGVHKSFRQVTTLSLQIPMGQFLKHLQSDASALR 780  
721 NMRHSESTYCIRQYAVVRDSDQGVHKSFRQVTTLSLQIPMGQFLKHLQSDASALR 780

QY 781 NSVIEQISINWESSSLFDFFLHFLRHSVVKIGDRCTOCQIGPOGSSILTLCLSCFG 840  
DB 781 NSVIEQISINWESSSLFDFFLHFLRHSVVKIGDRCTOCQIGPOGSSILTLCLSCFG 840  
QY 841 DMENKLFPAEVQDGLLRFVDDFLLVTPHLDQAKTFLSTLVHGVPEYGCMINLQKTVNMF 900  
DB 841 DMENKLFPAEVQDGLLRFVDDFLLVTPHLDQAKTFLSTLVHGVPEYGCMINLQKTVNMF 900  
QY 901 PVEPOTLGAAPYQLPAHCLPFWGCLLDLTQTLVFCDSYGAQTSIKTSLTFSQSVKAG 960  
DB 901 PVEPOTLGAAPYQLPAHCLPFWGCLLDLTQTLVFCDSYGAQTSIKTSLTFSQSVKAG 960  
QY 961 KTRMKLLSVLRKCHGLFLLQVNSLQTVCIINIKIFLLQAYRFHACVQLPFDQVRK 1020  
DB 961 KTRMKLLSVLRKCHGLFLLQVNSLQTVCIINIKIFLLQAYRFHACVQLPFDQVRK 1020  
QY 1021 NLTFELGISSQASCCYAILKVNPGMTLKASGFPPEAAHWLCYQAFLLKLAHSHVIYK 1080  
DB 1021 NLTFELGISSQASCCYAILKVNPGMTLKASGFPPEAAHWLCYQAFLLKLAHSHVIYK 1080  
QY 1081 CULGELRTAQKLLCRKPEATMTILKAAADPALSTDFQTILD 1122  
DB 1081 CULGELRTAQKLLCRKPEATMTILKAAADPALSTDFQTILD 1122

RESULT 2  
ABB06711  
ID ABB06711 standard; Protein; 1122 AA.  
XX ABB06711;  
XX  
XX 11-JUN-2002 (first entry)  
XX Mouse telomerase protein sequence.  
XX  
XX Mouse; telomerase; promoter; telomerase catalyst subunit; TERT; mTERT;  
XX enzyme; transgenic mouse; drug development; anticancer.  
XX Mus sp.  
XX JP2002000121-A.  
XX  
XX 08-JAN-2002.  
XX  
XX 23-JUN-2000; 2000JP-0190137.  
XX 23-JUN-2000; 2000JP-0190137.  
XX (RIKO-) ZH RIKOGAKU SHINKOKAI.  
XX (KIRI) KIRIN BREWERY KK.  
XX WPI; 2002-298279/34.  
XX  
XX A transgenic mouse comprising a DNA promoter region of mouse telomerase catalyst subunit (TERT) is used for the development of drugs and anticancer agents for regeneration of tissues and organs -  
XX Disclosure; Fig 3; 13pp; Japanese.  
XX  
XX The present invention describes a transgenic mouse (I) comprising a DNA construct having a DNA containing a promoter region of mouse telomerase catalyst subunit (TERT) and a DNA containing a reporter gene connected under the control of the promoter region. The transgenic mouse can be used in the development of drugs and anticancer agents for regeneration of tissues and organs. The present sequence represents the mouse telomerase protein, which is given in the exemplification of the present invention.

Query Match 99.2%; Score 5854; DB 23; Length 1122;  
Best Local Similarity 99.4%; Pred. No. 0;  
Sequence 1122 AA;

atches 1115; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 MTRAPCPAVRSLLRSRYREWVPLATFVRRLGEGRLVQPGDKPIYRTLVAQCUCVWHW 60  
1 MTRAPCPAVRSLLRSRYREWVPLATFVRRLGEGRLVQPGDKPIYRTLVAQCUCVWHW 60

61 GSQPPPADLSFHQVSSLKELVARVQORLCERNERNVLAFGFELLNARGGPPMAFTSSVR 120  
61 GSQPPPADLSFHQVSSLKELVARVQORLCERNERNVLAFGFELLNARGGPPMAFTSSVR 120

121 SYLPNTVETLRVSGAMWLLSRLVGGDDLLVYLAHCAVLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPNTVETLRVSGAMWLLSRLVGGDDLLVYLAHCAVLLVPPSCAYQVCGSPLYQICA 180

181 TTDIWPVSASVYAPTRPVGNFNLFLQIKSSRQCEAPKPLAPSRGCTKXHLSTSTS 240  
181 TTDIWPVSASVYAPTRPVGNFNLFLQIKSSRQCEAPKPLAPSRGCTKXHLSTSTS 240

241 VPSAKKARCYFVPRVEGPHRQVLPPTSGKSWVPSPARSPVPTAEKDLSSKGKVDLSL 300  
241 VPSAKKARCYFVPRVEGPHRQVLPPTSGKSWVPSPARSPVPTAEKDLSSKGKVDLSL 300

301 SGSVCCCHKSSSTSLSPRONAFQIRPPIETRHFLYSRGDQGERLNSFLSNLPNLT 360  
301 SGSVCCCHKSSSTSLSPRONAFQIRPPIETRHFLYSRGDQGERLNSFLSNLPNLT 360

361 GARLVEIIFLGSRPRTSGPLCRTHLSRYWQMRPLFQOLLVNHAECCOYVRLLSHCRF 420  
361 GARLVEIIFLGSRPRTSGPLCRTHLSRYWQMRPLFQOLLVNHAECCOYVRLLSHCRF 420

421 RTANQOVTDALNTPPHLMOLLHSPWQVYGLRACICKVVSASLWGTNRNRRFFKN 480  
421 RTANQOVTDALNTPPHLMOLLHSPWQVYGLRACICKVVSASLWGTNRNRRFFKN 480

481 LKRFISLGKVKLSLQELMKMKVEDCHWLRSRSGKDRVPAAEHRLRERILATFLFLMD 540  
481 LKRFISLGKVKLSLQELMKMKVEDCHWLRSRSGKDRVPAAEHRLRERILATFLFLMD 540

541 TYVVOLLRSFFYITESTFQKRLFFYKRSWMSKLSQIGVROHLRVLRLSQQEVRHHQ 600  
541 TYVVOLLRSFFYITESTFQKRLFFYKRSWMSKLSQIGVROHLRVLRLSQQEVRHHQ 600

601 DTWLAMPICLRFPKPNGLRPVNNVSYSGTALGRKQAOHFTQRLKTLFSLMNYERT 660  
601 DTWLAMPICLRFPKPNGLRPVNNVSYSGTALGRKQAOHFTQRLKTLFSLMNYERT 660

661 KPHLMGSSVLGWNIDYRTWRAFLVRLALDQTPRMVFKADVGTGAYDAIPOGKLVVVA 720  
661 KPHLMGSSVLGWNIDYRTWRAFLVRLALDQTPRMVFKADVGTGAYDAIPOGKLVVVA 720

721 NMIRHSESTYCIROYAVRRDSQGVHKSFRQVTTLSLQPMGQFLKHLQSDASALR 780  
721 NMIRHSESTYCIROYAVRRDSQGVHKSFRQVTTLSLQPMGQFLKHLQSDASALR 780

781 NSVVBOSISNNESSSIFDFFHLRHSVVKIGDRCYTCCQIGPGSSLSLTLCSLFG 840  
781 NSVVBOSISNNESSSIFDFFHLRHSVVKIGDRCYTCCQIGPGSSLSLTLCSLFG 840

841 DMENKLPFAVQORDGLLRFDVDFLLVTPHLDQAKTFLSTLVHGVPGYGCWMLQKTVNF 900  
841 DMENKLPFAVQORDGLLRFDVDFLLVTPHLDQAKTFLSTLVHGVPGYGCWMLQKTVNF 900

901 PVEPGLTGGAPYOLPAHCLFPKCGLLDTQTLEVCDSYGAQTSIKTSLNFQSVFKAG 960  
901 PVEPGLTGGAPYOLPAHCLFPKCGLLDTQTLEVCDSYGAQTSIKTSLNFQSVFKAG 960

961 KTMNKLLSVLRKCHGLFLDLQVNSLQTYCINIYKIFLLQAVRPHACVQLPFDQVRK 1020  
961 KTMNKLLSVLRKCHGLFLDLQVNSLQTYCINIYKIFLLQAVRPHACVQLPFDQVRK 1020

1021 NLTFPLGIISSQASCCVAILKVRNPGMTLKASGSPPEAAHMLCYQAFLLKLAHSVYIK 1080  
1021 NLTFPLGIISSQASCCVAILKVRNPGMTLKASGSPPEAAHMLCYQAFLLKLAHSVYIK 1080

Qy 1081 CLIGPLRTRAQKLLCKLPEATWTLKAAADPALSTDPQTILD 1122  
Db 1081 CLIGPLRTRAQKLLCKLPEATWTLKAAADPALSTDPQTILD 1122

## RESULT 3

AAW46957  
ID AAW46957 standard; Protein; 1132 AA.

XX AC AAW46957;

XX DT 13-AUG-1998 (first entry)

XX Human telomerase reverse transcriptase.

XX Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis;  
KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.

XX OS Homo sapiens.

XX GN GB2317891-A.

XX PD 08-APR-1998.

XX PF 01-OCT-1997; 97GB-0020890.

XX PR 14-AUG-1997; 97US-0915503.

XX PR 01-OCT-1996; 96US-0724643.

XX PR 18-APR-1997; 97US-0844419.

XX PR 25-APR-1997; 97US-0846017.

XX PR 06-MAY-1997; 97US-0851843.

XX PR 09-MAY-1997; 97US-0854050.

XX PR 14-AUG-1997; 97US-0911312.

XX PR 14-AUG-1997; 97US-0912951.

XX (GERO-) GERON CORP.

XX (UYTE-) UNIV TECHNOLOGY CORP.

XX Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J;  
PI Morin GB, Nakamura T, Harley CB;

XX WPI; 1998-171633/16.

XX N-PSDB; AAV22379.

XX Pure and recombinant human Telomerase Reverse Transcriptase and its  
PT variants - are useful in the diagnosis, prognosis and treatment of  
PT cell proliferation conditions especially cancer and ageing

XX Claim 3; Fig 17; 387pp; English.

XX The present sequence represents human telomerase reverse transcriptase  
CC (hTERT), which is a ribonucleoprotein. The present invention also  
CC describes the following methods: (A) determining whether a test compound  
CC is a modulator of hTERT, by detecting the change in hTERT recombinant  
CC protein or polynucleotide, on administration of the compound; (B)  
CC preparation of recombinant telomerase by contacting a protein  
CC preparation of hTERT with a telomerase RNA component; (C) detection of  
CC the hTERT RNA or protein in a sample by binding a relevant probe to the  
CC sample and detecting the complex formed or in the case of RNA detection,  
CC amplifying the product and correlating the presence of complex or  
CC amplification product with presence of hTERT in the sample; and (D)  
CC increasing the proliferation of a vertebrate cell by increasing hTERT  
CC expression; and (E) the use of an agent that causes an increase in cell  
CC vertebrate cell proliferation to create a medicament that inhibits  
CC ageing. A protein preparation of hTERT and the polynucleotide encoding  
CC hTERT can be used in the manufacture of medicaments for inhibiting the  
CC effect of ageing or cancer. Inhibitors of telomerase activity can be  
CC used to treat conditions that are associated with high telomerase  
CC activity. A protein preparation of hTERT can also be used in the new  
CC methods.

XX Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 19; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLLSRVREVMPLATFVRRIGPEGRRLVQDGPDKIYETILVAOCLVCMHW 60  
1 MTRAPCPAVRSLLSRVREVMPLATFVRRIGPEGRRLVQDGPDKIYETILVAOCLVCMHW 60  
61 GSOPPPADLSHFQVSSKELVARVVQRLCERNERNVLAFFGELLNEARGPPNAPTSSVR 120  
61 DAPPPAPAFSPQVSCVKELVARVQLRCERGAKNVLAFFGALLDGGARGPPPEAFTTSVR 120  
121 SYLPTNVTIETRVSGAWMLLSRVGDDLLVILLAHCALYLLVPSCAYOYCGSPLIQICA 180  
121 SYLPTNVTIETRVSGAWMLLSRVGDDLLVILLAHCALYLLVPSCAYOYCGSPLIQICA 180  
181 TTDIWPSVASYPTRPVGRNTLRFLLQIKSSSQEQEAPKPLALPSGRGTRKHLSTSTS 240  
181 ATQAREPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAPGARRRGSASRS 231  
241 VPSAKARCYVPVREVEGP-----HRQVLTPTSGKSW-VPSPARSEVPTAKDLSSK 292  
232 LPLPKPRRGAAPERTPVGSGWAHPGTRGSDRGFCVVSAPAR-----PABEATSLE 286  
293 GKVSIDLSS-GSVCCCHKPSSLSLPPRONAFQLRP-PIETRHLYSGDGOERLNPSP 350  
287 GALSCTRHSHPSVGRQHAGAPSTSEPRPMDTPCPVYAKTKHLYSSGD-KQOLRPSF 345  
351 LLSNLQPNLTGAARLVEIIFLGSRRPTSGPLCORTHLRRYQWQMRPLFQQLLVNHAECQY 410  
346 LLSLRPLSLTGARLVETIFLGSRRPMDTPRRLPRLPQRYWQMRPLFLELLGNHACQY 405  
411 VRLRSHCFRTANQOVTAL-----NTSPHMLDLRLHSSPQVY 452  
406 GVLLKTHCPURAA---VTPAAGVCAKEPQGVAAPEBEDTPRRLVOLLQKQSSPQVY 462  
453 GFLRACLKVVVSASLMTGRHNRERRFKLKFISLQKYLQSLQELMMKXVEDCHLRS 512  
463 GFVRACLRRLVPLPGLWSRHNERRFLNFKFISLQKYLQSLQELMMKXVEDCHLRS 522  
513 SPGRDVPAAHRLRRLIATLFWLMDYVYVQLLRSFFYITESTFQKRLFFRKSVWS 572  
523 SPGVGCVPAEAEHLREIEILAKELHMLMSYVYVELLRSFFYITESTFQKRLFFRKSVWS 582  
573 KLOSIGVROHLERVLRELISOEVRHODTLAMPICRLRFPKPNGLRPVWNSYSGMT 632  
583 KLOSIGVROHLERVLRELISOEVRHODTLAMPICRLRFPKPNGLRPVWNSYSGMT 642  
633 RALGRKQAQHTORLKTFLSMNLVYRTKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
643 RTFRREKRAEHLTSRVKALFVSLNVERARRPGLLGASVLGLDDIHRARWTFVLRVRAQDP 702  
693 TPBMVYKADVTGAYDAIPQGLKLVAVVANNIRHSESTYCIROYAVVRSDGOVHKSPER 752  
703 PPELYFVKVDVTGAYDTIPDRUTEVIAIIG-KPNTYCVRIYAVVQAAHGHVKAFAKS 761  
753 QVTLSDLPQYMGQFLKHLQSDASALRNSVVEIOSIMNESSSSLFDFLHLRHSVYK 812  
762 KVSTLTLQPYMGQFAHLQET--SPLRDAVVEIQSSSLNEASSGLFDVFLRFMCHHAVR 819  
813 IGDECYTQCOGIPQSSSLTLCSLCFQGMENKLPRAVQDGLLRFVDDLLVTPHLQD 872  
820 IRGKSYVCCQGIPOGSLTSLTLCSLCYGMENKLPAGIRRDGLLRLVDDFLVTPHLTH 879  
873 AKTELSTLVHGVPEYGMINLQKTVNVPFVEPOTLGGAAFYQLPAHCLFPWCGLLDQTQ 932  
880 AKTELSTLVHGVPEYGMINLQKTVNVPFVEPOTLGGAAFYQLPAHCLFPWCGLLDQTQ 939  
933 LEVCDYSGVAQTSIKTSLTFQSFVAGKTMKNKLSVLRLKCHGLFLDLQVNSLQTVCI 992  
940 LEVQDSYSSYARTSIRASLTFNFGKAGNMRKLFGLVRLKCHSLFLDLQVNSLQTVCI 999

QY 993 NIYKIFLOAYRHACVIOQLPFDQRYKRLTFFLGLIISQASCCYAILKVKPQMTLKAS 1052  
DB 1000 NIYKIFLOAYRHACVIOQLPFDQRYKRLTFFLGLIISQASCCYAILKVKPQMTLKAS 1059  
QY 1053 GS---FPPEAAHMLCYQAFLLKLAASHVYKLLGLPLRTAQKLLCKLPEATMTILKAAA 1109  
DB 1060 GAAGPLSEAVQWMLCHQAFLLKLTNRVTVVPLGLSRTAQKLLCKLPEATMTILKAAA 1119  
QY 1110 DPALSTDFQITLD 1122  
DB 1120 NPALPSDFKITLD 1132

RESULT 4  
AA43621  
ID AAY43621 standard; Protein; 1132 AA.  
XX AC AAY43621;  
XX DT 26-JAN-2000 (first entry)  
XX DE A human telomerase reverse transcriptase (TRT) polypeptide.  
XX KW Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;  
XX KW dendritic cell; telomerase activity; cancer cell; proliferating cell;  
XX KW immunological destruction; telomerase; cancer; proliferation disease.  
XX OS Homo sapiens.  
XX FN WO950392-A1.  
XX PD 07-OCT-1999.  
XX PF 30-MAR-1999; 93WO-US06898.  
XX PR 31-MAR-1998; 98US-0112006.  
XX (GERO-) GERON CORP.  
XX Gaeta FCA;  
XX WPI: 1999-610845/52.  
XX N-PSDB; AAZ30154.  
XX ELICITING an in vivo immune response for prevention and treatment of  
PT cancers  
PS Claim 3; Fig 1; 26pp; English.  
XX  
CC The present sequence represents a human telomerase reverse transcriptase  
CC (TRT) polypeptide. The protein is used in the method of the invention.  
CC The specification describes a method for activating a T lymphocyte,  
CC comprising contacting the T lymphocyte with a dendritic cell that  
CC expresses a TRT peptide in the context of a MHC class I or MHC class  
CC II molecule. The protein causes induction of an in vivo immunological  
CC response to telomerase activity. Cancer cells are characterized by  
CC expression of endogenous trt gene and the presence of detectable  
CC telomerase activity. Therefore, by eliciting a specific immune response  
CC to TRT or to TRT-expressing cells, it is possible to selectively target  
CC proliferating cells for immunological destruction. The method is used  
CC for eliciting an in vivo immune response to telomerase by activating  
CC a T lymphocyte, and is useful for prevention and treatment of cancers and  
CC other proliferation diseases/conditions.  
XX SQ Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 20; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLLSRVREVMPLATFVRRIGPEGRRLVQDGPDKIYETILVAOCLVCMHW 60  
1 MTRAPCPAVRSLLSRVREVMPLATFVRRIGPEGRRLVQDGPDKIYETILVAOCLVCMHW 60

[illegible]

```

121 SYLNTVETLRVSGAMWLLLSRVGDDLLVYLLAHCALYLLVPPSCAYOVCSPLYQICA 180
121 SYLNTVTDALRGSCAMGLLRLRVGDDVLVHLLARCALFVLVAPSAYQVCPPLYQLGA 180
181 TTDIWPVSASRYRTFVGNFTNLRFLQIKSSRQEAQPKLALSRGTRKRLSLTSTS 240
181 ATQARPPHAS-GPRLRLG-----CERAWHSVREACVPLGLPAPGARRRGSASRS 231
241 VPSAKARCYVPVRVERGP-----HROVLPSPGKSM-VSPAPSPVPTAEXDLISK 292
232 LPLPKRPRGAAPERTFVQGSNAHPRGTRGFSDRGFCVWSPAR-----PAEASISLE 286
293 GKVSPLSLG-OSVCKHKPSSTLSLSPRQNAQLRP-FLETRHFLYSRGDSOERLNSP 350
287 GALSGRTHSHSVGRQHAGPPSTSPRPWDTPCPVVAETHKFLYSSGD-KEQRSP 345
351 LLSNLQPNLTGARRLVRIIFLGSRPRTSGCLRTHLSRRYQMRFLFQCLLVNHAECQY 410
346 LLSLRPSLTGARRLVETIFLGSRPWMPGTFRPLPLPQRYWQMRPLFLELLGNHAQCFY 405
411 VRLLSHCRFTANQOVTDAL-----NTSPHLMDLRLHSSPQVY 452
406 GVLKTHCPLRAA---VTPAAGVAREKPGSVAAPEEEDTPRRLVQLRHSPPQVY 462
453 GFLRACLCKVVSASLWGTNRNERRFFKNLKKFISLKGKLSLQELMWKVKVEDCHWLS 512
463 GFVRACLRLVPPGLWGSRNERRFLNTKKFISLKGKAKLSLQELTWKSVRDCAWLRR 522
513 SPGKDRVPAABHRLRERILATFLFWIMDMTVVQLLRSPFFYITESTQKRLFPYKSVWS 572
523 SPFGVCVPAABHRLREELAKFLHMLS VYVELLRSPFFYITESTQKRLFPYKSVWS 582
573 KLOSIGVQHLEVRVRLRELSOEVRHDDTWLAMPICRLRFIPKPNGLRPIVNMVSMGT 632
593 KLOSIGVQHLEVRVRLRELSOEVRHDDTWLAMPICRLRFIPKPNGLRPIVNMVSMGT 642
633 RALGRKQAQHTORLKTFLFNLNYERTKPHLMGSSVLGMNDIYTWRAFVLRVRLDQ 692
643 RTFREKRAELTSRVKALFVNLVYERARRPGLLGASVLGLDIDHRAWRTFVLRVAQDP 702
693 PRMYFVKADVTGADATPOCKLVEVNVANIRHSENYC-RQVAVVRDSQGOVHKSPPR 752
703 PDELTVFVVDVTGADVTTPQDLRLTEVIAELIK-PQNYCVRRVAVVQKAAHGVRAFKS 761
753 QVTTLSLQPMYQGPFLKHLQSDASALNSVYIEQSISNMNESSSLFDFPLHFLRHSVYK 812
762 HVSTLTDLPYMRQFVAHQET--SPLRDAVVIQSSSLNEASSGLFDVFLRPMCHAVR 819
813 IDRCYTCOGIPQOSSLSTILCSLFCGDMENKLPFAVQSDGILLRFPVDDFLVTLHLDQ 872
820 IRGKSYVQCOGIPQSSILSTLLCSLCYGD MENKLPFAGIRDRGLLRVDDFLVTLHLDQ 879
873 AKTFSLTLVHGVPEYCGMINLQKTVNPPPEPGLGGAAPYQLPAHCLFPWCGLLLDQ 932
880 AKTFSLTLVHGVPEYCGVNLKTVNPPPEPGLGGAAPYQLPAHCLFPWCGLLLDQ 939
933 LEVFCDSYGAQTSITKSLTQSVFKAGTKWRNKLKSLVRLKHKGLFLDQVNSLQTVCI 992
940 LEVQSDYSYARTSIRASLTNROGFKAGNNRKLFGVLRKCHKSLFLDQVNSLQTVCT 999
993 NYKIFLLQAVRFHACVTLQLPDQVRNQLTFFLGIISQASCCYAILKVKNPMTLKAS 1052
1000 NYKILLQAVRFHACVTLQLPDQVRNQLTFFLGIISQASCCYAILKVKNPMTLKAS 1059
1053 GS---FPPEAHWLCYQAFLLKLAHSHVYIKLGLPLRTAQKLLCRKLPEATWTILKAA 1109
1060 GAAGLPSEAVQWLCHQAFLLKTHRVYTVYVLLGSLRTAQTLQSLRKLPGTTLTALEAAA 1119
1110 DPALSTDFQTLID 1122
1120 NPALPSDFKTLID 1132

```

```

RESULT 6
AAY32090
ID AAY32090 standard; Protein; 1132 AA.
XX
AC AAY32090;
XX
DT 17-JAN-2000 (first entry)
XX
DE Human telomerase reverse transcriptase (hTERT).
XX
DE Telomerase reverse transcriptase; human; hTERT; cell proliferation;
XX
KW cancer.
XX
OS Homo sapiens.
XX
XX W09950386-A2.
XX
XX 07-OCT-1999.
XX
XX 31-MAR-1999; 99WO-US07097.
XX
XX 31-MAR-1998; 98US-0052864.
XX
XX 03-AUG-1998; 98US-0128354.
XX
XX (GERO-) GERON CORP.
XX
XX Morin GB;
XX
XX WPI; 1999-610842/52.
XX
XX N-PSDB; AAZ20279.
XX
XX New catalytic polypeptide and polynucleotide, useful for increasing
XX
XX catalytic activity in a cell -
XX
XX Claim 13; Fig 1; 24pp; English.
XX
XX The present sequence represents human telomerase reverse
XX
XX transcriptase (hTERT). Human telomerase is a target for diagnosing
XX
XX and treating diseases relating to cell proliferation and
XX
XX senescence, such as cancer, or for increasing the proliferative
XX
XX capacity of a cell. A claimed method for increasing the
XX
XX proliferative capacity of a vertebrate cell, especially a human or
XX
XX other mammalian cell, involves introducing into the cell a
XX
XX recombinant hTERT polynucleotide encoding an hTERT variant in which
XX
XX residues 192-323, 200-323, 192-271, 200-271, 222-240, 415-450,
XX
XX 192-323 and 415-450, or 192-271 and 415-450 of the present sequence
XX
XX are deleted. A claimed method of preparing recombinant telomerase
XX
XX involves contacting a recombinant hTERT deletion mutant (as above)
XX
XX with a telomerase RNA component such that the 2 proteins associate
XX
XX to form a complex capable of catalysing the addition of nucleotides
XX
XX to a telomerase substrate. A claimed method for reducing
XX
XX telomerase activity in a cell involves introducing a recombinant
XX
XX polynucleotide encoding an hTERT variant having a deletion of amino
XX
XX acids 192-450, 560-565, 637-660, 638-660, 748-764 or 1055-1071 of
XX
XX the present sequence.
XX
XX Sequence 1132 AA;
XX
XX Query Match 59.4%; Score 3505; DB 20; Length 1132;
XX
XX Best Local Similarity 62.4%; Pred. No. 0;
XX
XX Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;
XX
XX
XX 1 MTRAPRCVARSLLRSRYREVWPLATFVRRLQPEGRRLVQPDGPKIYRTLVAAQCLVCMHW 60
XX
XX 1 MPRAPRCVARSLLRSRYREVWPLATFVRRLQPEGRRLVQPDGPKIYRTLVAAQCLVCMHW 60
XX
XX 61 GSQPPPADLSFHQVSSKELVARVQVRLCERNERNVLAFGFELLNEARGPPMATTSVR 120
XX
XX 61 DARPPPAAPSRFQVSCLEKELVARVQVRLCERNERNVLAFGFELLNEARGPPMATTSVR 120
XX
XX 121 SYLNTVETLRVSGAMWLLLSRVGDDLLVYLLAHCALYLLVPPSCAYOVCSPLYQICA 180

```



```

121 SYLNTVTDALRGSGAWGLLRRVGDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
181 TTIDWPSVASYRTPVGRNFTNLRFLOQIKSSROEAPKPLAPSGRTKRLSLTSTS 240
181 ATQARPPPHAS-GPERRLG-----CERAWNHSVREAGVPLGLPAGHARRGGASRS 231
241 VPSAKKACVYPRVVEGP-----HRQVLPYSGKSW-VPSPARSPVPTAEKDLSSK 292
232 LPLKPRPRGAPEPRTFVQGGSWAHFGRTRGSDRGFCVSPAR-----PAEATSLE 286
293 GKVSLSLS-GSVCKHKPSSTLSLSPRONAFQLRP-FIETRHFLYSGDGOERLNPSF 350
287 GALSSTRHSHSVGRQHAGPSTSRPPRWDTPCPVYAEKHFLYSSGD-KEQLRPSF 345
351 LLSNLQPNLTGARRLVEIFLGSRRPSTGCLTRHLSRRYQWRPLFOQLLVNHAQQY 410
346 LLSLRPLSLTGARRLVEIFLGSRRPSTGCLTRHLSRRYQWRPLFOQLLVNHAQQY 405
411 VLLRSHCFRTANQVTDAL-----NTSPHMLDLRLHSSPWQVY 452
406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRLVQLLRQHSSPWQVY 462
453 GFRLACLCVKSASLWGRHNERFFKNLKKFISLGKYGKLSLQELMWKMKVDCEHLRS 512
463 GFVRACLRLVPLGLWGRHNERFFKNLKKFISLGKXKLSLQELTWKMSVRDCAWLR 522
513 SPGKDRVPAAEHRLRERILATFLWMDTYVQLLRFFYITESTFOKNLFPYKSVMS 572
523 SPGVCVPAAEHRLRERILATFLWMDTYVQLLRFFYITESTFOKNLFPYKSVMS 582
573 KLOSGVGRHRLRERILATFLWMDTYVQLLRFFYITESTFOKNLFPYKSVMS 632
583 KLOSGVGRHRLRERILATFLWMDTYVQLLRFFYITESTFOKNLFPYKSVMS 642
633 RALGRKQAHQTBQRLKTLFSLMNYRTKPHLMGSSVLGMNDIYRWTRAPVLRALDQ 692
643 RTFREKGAERLTSVKALFSLVNERARRPGLLGASVGLDDIHRAWRFTVLVRAQDP 702
693 TRPVFVADVTGAYDAIPOGKLVEVAVNMIRHSESTVCIROAVVRDQGOVHKSFR 752
703 PELFVVRVDVTGAYDTIPQRLTEVIAIIT-KONTYCVRYAVVQKAAHGVRAKFS 761
753 QVTTLSLDLPYMGQFLKHLQSDASALRNSVVIQSISMNESSSLDFLHLRHSVX 812
762 HVSTLTDLQPYMRQFVAHLQET--SPLRADAVVIRQSSSLNEASSGLFDVFLRFMCHAVR 819
813 IGRCYTQCGIPQSSSLTLCSCFGDMENKLPFAEVORDGILLRVDFFLLVTHLDQ 872
820 IRKSVYQCGIPQSSSLTLCSCFGDMENKLPFAEVORDGILLRVDFFLLVTHLDQ 879
873 AKTFLLTVHGVPEYGCMLNLOKTVNPPVPGTILGGAAPYQLPAHCLFPWCGLLDTQT 932
880 AKTFLLTVHGVPEYGCMLNLOKTVNPPVPGTILGGAAPYQLPAHCLFPWCGLLDTQT 939
933 LEVFDYGYAOTSITSITPSVFKAGTWRNKLSSVLRLKCHGLFLDLQVNSLQVCI 992
940 LEVQSDYSYARTSIRASLTNRGFKAGNWRKLFGLVRLKCHSLFLDLQVNSLQVCI 999
993 NYIKFLQAYRPHACVQLPQDQVRKVLNLTFFLGISSQASCCVAILKVNKPGMTLKAS 1052
1000 NYIKFLQAYRPHACVQLPQDQVRKVLNLTFFLGISSQASCCVAILKVNKPGMTLKAS 1059
1053 GS---PPPAHWCYQAFLLKLAHSHVYIKCLGLPLTAQKLCKRLPEATMTLKAAA 1109
1060 GAAGPLPSEAVQWCHQAFLLKLTSHRVTVYVPLLSLRTAQQLSRKLPGLTTLAEAAA 1119
1110 DPALSTDFOTILD 1122
1120 NPALPSDFKTIILD 1132

```

AY26580 standard; Protein; 1132 AA.  
AY26580;  
13-SEP-1999 (first entry)  
Human telomerase reverse transcriptase (hTERT) enzyme.  
Telomerase reverse transcriptase; TERT; mouse; telomere length assay;  
immunogen; enzyme; telomerase-mediated DNA replication; human.  
Homo sapiens.  
WO9927113-A1.  
03-JUN-1999.  
25-NOV-1998; 98WO-US25211.  
16-MAR-1998; 98US-0042460.  
26-NOV-1997; 97US-0979742.  
(GERO-) GERON CORP.  
(YESH) UNIV YESHIVA EINSTEIN COLLEGE.  
Allsopp R, Depinho R, Greenberg R, Morin GB;  
WPI; 1999-347722/29.  
Mouse telomerase reverse transcriptase (mTERT) enzyme proteins and  
nucleic acids  
Disclosure; Fig 3; 135pp; English.  
The invention relates to a mouse telomerase reverse transcriptase (mTERT)  
enzyme. Compositions containing mTERT can be used in telomere length  
assays. Isolated mTERT is useful as an immunogen for the production of  
monoclonal or polyclonal antibodies. The method is useful for assessing  
the degree of purification and identification of new mTERT species, such  
as an mTERT allele, homolog or isoform, or to screen for modulators  
(antagonists and agonists) of telomerase-mediated DNA replication.  
Antagonists and agonists of mTERT can be used to modify the activity of  
other telomerase enzymes such as human TERT (hTERT). The present sequence  
represents a human TERT enzyme.  
Sequence 1132 AA;  
Query Match 59.4%; Score 3505; DB 20; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPCPAVRSILRSRYREVWPLATFVRRLGPEGRRLVQPGDPKTYRTLVAQCLVCMHW 60  
Db 1 MPAPRCVAVRSILRSRYREVWPLATFVRRLGPEGRRLVQPGDPKTYRTLVAQCLVCMHW 60  
QY 61 GSQPPADLSHQVSSIKELVARVVOQLCERNERNVLAFGELLNEARGGPPMAFTSSVR 120  
Db 61 DARPPPAAPSPQVSCIKELVARVVOQLCERNERNVLAFGELLNEARGGPPMAFTSSVR 120  
QY 121 SYLNTVTETLRVSGAMWLLSRVGGDLLVYLLAHACALYLLVPPSCAYQVCGSPLYQICA 180  
Db 121 SYLNTVTETLRVSGAMWLLSRVGGDLLVYLLAHACALYLLVPPSCAYQVCGSPLYQICA 180  
QY 181 TTIDWPSVASYRTPVGRNFTNLRFLOQIKSSROEAPKPLAPSGRTKRLSLTSTS 240  
Db 181 ATQARPPPHAS-GPERRLG-----CERAWNHSVREAGVPLGLPAGHARRGGASRS 231  
QY 241 VPSAKKACVYPRVVEGP-----HRQVLPYSGKSW-VPSPARSPVPTAEKDLSSK 292  
Db 232 LPLKPRPRGAPEPRTFVQGGSWAHFGRTRGSDRGFCVSPAR-----PAEATSLE 286  
QY 293 GKVSLSLS-GSVCKHKPSSTLSLSPRONAFQLRP-FIETRHFLYSGDGOERLNPSF 350

287 GALSCTRSHSPVSGRQHAGPPTSRPRPWRDTPCPVPVYAEYKHYLYSSGD-KEQLRPSF 345  
 351 LLSNLQPNLTGARRLVETIIIFLGSRPTSGPLCTHRLSRRYWQMRPLFQQLVNHAEQY 410  
 346 LLSLRPSLTGARRLVETIIIFLGSRPMPGTPRLPRLPQRYWQMRPLFELLGNAQCFY 405  
 411 VRLASHCRFTANQOVTDAL-----NTSPHLMDLRLHSSPWQY 452  
 406 GYLLKTHCPLEAA---VTPAGVCAREKPGQSVAAPEEDTDPRLVQLLRHSSPWQY 462  
 453 GFLRACLCKVVSASILWGRHNERFFKNLKKFISLQKYGKLSLQELMWMKVEDDCHWLR 512  
 463 CFVRACLRLVPPGLMGRHNERFLENTKFIISLQKHAKLSLQELTWMSVEDCAWLR 522  
 513 SPGKDRVPAAEHRLRERILATFLFWIMDMTVVQLLRSPFYITESTQKRLFPYKSVWS 572  
 523 SPGVGCPAAEHRLREELAKFLHLMVSVVVELLSFFYVTTTQKRLFPYKSVWS 582  
 573 KLOSTGVQHLERVLRLRELSOEYRHHQDTWLAMPICLRLFIKPNGLRPIVNMYSMT 632  
 583 KLOSTIGIQLHRLVQLRELSAEVQCHREARPAALLTSRLRFIPKPLGRPIVNMVYGA 642  
 633 RALGRKQAOHFTQRLKTLFSLMNYERTKPHLMGSSVLGMNDIYRTWAFVLRVADQ 692  
 643 RTFREKRAELTGRVKALFSLVNYERARRPGLLGASVLGLDDIHEARWTFVLRVAQDP 702  
 693 TPRMYFVKADVTGAYDAIPOGKLVVVANMRHSESTYCIQVAVVRDQSGOVHKSFR 752  
 703 PPELVFVKVDVTGAYDTIIPQRLTEVIASIIK-PQNTYCVRRYAVVQKAHGVKAFKS 761  
 753 QVTTLSLDLPQYMGFLKHLQSDASALRNSVYIEQSIQSNNESSSLFDFPLHFRHSVVK 812  
 762 HYSTLTDLQPYMRQFVAHQET--SPLEDAVIEQSSSLNEASSGLFDVFLRPMCHAVR 819  
 813 IDRCYTCQGIPOGSSUSTLLCSLFCGDMENKLPFAVQRDGLLRFVDFVFLVTHLQ 872  
 820 IRGKSYVQCSIPQSSILLTLLCSLGYDMENKLPFAGIRDRGLLRLVDFLLVTHLTH 879  
 873 AKTFTSLTVHGVPEYGCINLQKTVNPPVPPGTLGGAAPVQLPAHCLFPWCGLLDDTQT 932  
 880 AKTFTLVLRGVPEYGCVNVNLRKTVNPFVEDEALGCTAFVQVPAHCLFPWCGLLDDT 939  
 933 LEVFCDSYGAQTSIKTSLTQSVFKAGTKWRNKLVLRLKCHGLEFLDQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASILTNRGFKAGNRMRKLFGLRLKCHSLFELDQVNSLQTVCT 999  
 993 NIYKIFLLOAYRFHACVQLPDPQVRKNTLFFELGISSQASCCYAILKVNPMWTLKAS 1052  
 1000 NIYKILLQAYRFHACVQLPDPHQVKNPTFFLRVSDTASLSCYSLKAKNAGNSLGAK 1059  
 1053 GS---FPPAAHWCYQAPLLKLAHSHVYIKLLGLPLRTAQKLCRLKPLATMTILKAAA 1109  
 1060 GAAGPLPSEAVQWLCHQAPLLKLTSHRYTVYVPLGLSLRTAQQLSRKLPGLTTLALEAAA 1119  
 1110 DPALSTDFQTIID 1122  
 1120 NPALPSPDFXTIID 1132

SULT 8  
 W90251  
 AAW90251 standard; Protein; 1132 AA.  
 AAW90251;

24-MAY-1999 (first entry)

Human catalytic telomerase sub-unit protein.

Human; catalytic telomerase subunit; therapy; diagnosis; htc; assay; modulator; treatment; inhibit; cellular disorder; death; defect; cancer; ageing; antisenase; neoplastic cell; telomerase-related condition; tumour cell.

XX OS Homo sapiens.  
 XX FN WO9859040-A2.  
 XX PD 30-DEC-1998.  
 XX PF 09-JUN-1998; 98WO-EP03468.  
 XX PR 14-APR-1998; 98DE-1016496.  
 XX PR 20-JUN-1997; 97DE-1026329.  
 XX PR 26-MAR-1998; 98DE-1013274.  
 XX PA (FARB ) BAYER AG.  
 XX Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;  
 PI WI; 1999-081276/07.  
 XX DR N-PSDB; AAV72117.  
 XX DR New catalytically active subunit of human telomerase - used in the  
 PT modulation of telomerase activity, particularly for treating cancer  
 PT and ageing  
 XX Claim 2; Fig 2; 76pp; German.  
 XX This sequence represents a novel human catalytic telomerase sub-unit  
 CC (htc). This protein can be used in screening assays to identify  
 CC modulators of telomerase and to treat or inhibit cellular disorders,  
 CC death, defects and/or other pathological processes involving telomerase,  
 CC particularly cancer and ageing (also suitable for this are agents that  
 CC stimulate, inhibit or mimic the activity of the subunit). Antisense  
 CC nucleic acids inhibit telomerase action (by binding to specific mRNA),  
 CC particularly in neoplastic cells and may be expressed in vivo. Antibodies  
 CC and fragments of the protein, used as probes or primers, are used to  
 CC diagnose telomerase-related conditions (especially neoplasia) by (i)  
 CC detecting abnormal levels of the subunit protein in body fluids or  
 CC tissues or (ii) by measuring the amount of the encoding nucleic acid.  
 CC Expression of the nucleic acid encoding the subunit mRNA is confined to  
 CC tumour cells, in contrast to the ubiquitous expression of the telomerase  
 CC RNA subunit.  
 XX SQ Sequence 1132 AA;  
 Query Match 59.4%; Score 3505; DB 20; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 QY 1 MTRAPCPAVRSLLRGRYREVWPLATFVRLGPEGRRLVQPDPKIYRTLVAQCLVCMHW 60  
 DB 1 MPAPRCRAVRSLLRSHYREVPLATFVRLGQGRVLRVQGPAPRALVAQCLVCPW 60  
 QY 61 GSQPPPADLSFHQVSSILKELVARVQRLCERNERNVLAFGFELLNEARGGPPMAFTSVR 120  
 DB 61 DARPPPAAPSFQVSCLELVARVQLCERGAKNVLAFGFALLDARGGPPPEAFTSVR 120  
 QY 121 SYLNTVIEILRVSGAWMLLSRVGDDLLVLAHACALVLLVPPSCAYOVCGSPYQICA 180  
 DB 121 SYLNTVTDALRGSGAWGLLRVGDVLLVHLLARCALFVLVAPSCAYOVCGFPPLYLGA 180  
 QY 181 TTDIWPVSASVYRPTREVNGENTNRLFLQIKSSSSRQEAAPKPLALPSRGTKRHLSTSTS 240  
 DB 181 ATQAPPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAGARRRGGSSASRS 231  
 QY 241 VPSAKARCPVPVREVEGP-----HRQVLTPPSGKSW-VSPARSPEVTAEKDLSK 292  
 DB 232 LPLPKRPRGAAPPEPFTFVGQGSWAHPGTRGSPDRGFCVVSAPR-----PAEATSLSE 286  
 QY 293 GKVSDDLSS-GSVCKKPSSTSLSPRONAFQLRP-FETKHFYLSRGDQGERLNPSF 350  
 DB 287 GALSCTRSHSPVSGRQHAGPPTSRPRPWRDTPCPVPVYAEYKHYLYSSGD-KEQLRPSF 345  
 QY 351 LLSNLQPNLTGARRLVETIIIFLGSRPTSGPLCTHRLSRRYWQMRPLFQQLVNHAEQY 410



15-FEB-2001; 2001WO-US05143.

15-FEB-2000; 2000US-0182685.

15-FEB-2001; 2001US-0182685.

(REGC ) UNIV CALIFORNIA.

Zanetti M;

WPI; 2001-536552/59.

Vaccine for initiating and enhancing a cytotoxic T lymphocyte response, for treating cancers or tumors or for inducing immune response against tumors, comprises a telomerase reverse transcriptase peptide

Disclosure; Fig 5; 52pp; English.

The present sequence is that of human telomerase reverse transcriptase (hTERT). The sequence was analysed for 9-mer peptide sequences containing known binding motifs for the human leukocyte antigen HLA-A2.1 molecule. From an initial panel of about 30 candidate peptides, 2 sequences, denoted p540 (see AAB82772) and p865 (see AAB82773), were examined. The majority of healthy individuals as well as patients with prostate cancer immunised in vitro against these 2 HLA-A2.1 restricted peptides developed hTERT-specific cytotoxic T lymphocytes (CTL). The cancer patients' CTL specifically lysed a variety of HLA-A2+ cancer cell lines such as prostate, breast, colon, lung and melanoma, demonstrating immunological recognition of endogenously-processed hTERT peptides. In vivo immunisation of HLA-A2.1 transgenic mice generated a specific CTL response against both hTERT peptides. The induction of CTL responses in vitro and in vivo, and the susceptibility to lysis of tumour cells of various origins by hTERT CTL suggest that hTERT could serve as a universal cancer vaccine for humans. Thus, a claimed universal vaccine for treating tumours of any origin comprises at least 1 hTERT peptide in an amount effective for initiating and enhancing a CTL response against cancer cells. The peptide is 7-15 amino acid residues in length and may be modified to enhance binding to the major histocompatibility complex. Also claimed is a method for inducing and enhancing a CTL response against cancer cells, involving harvesting blood leukocytes, pulsing with hTERT, and contacting cancer cells with the pulsed leukocytes. A method for targeting CTL to tumour cells is also claimed, and involves administering a hTERT peptide to a mammal, especially a cancer patient.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
Y 1 MTRAPRCVAVRSLRSRYREVWPLATVRLGPEGRLVQGDPKYRVLVAQCLVCMHW 60  
b 1 MERAPRCVAVRSLRSRYREVWPLATVRLGPEGRLVQGDPAFALVAQCLVCPW 60  
Y 61 GSQPPADLSFQVSSSLKELVARVQBLCEBENRVLAFGFEINAEARGPPMAFTSSVR 120  
b 61 DARPPPPAAPSPFQVSSSLKELVARVQBLCEBENRVLAFGFEINAEARGPPMAFTSSVR 120  
Y 121 SYLPTNVTETLRVSGAWMLLSRVGDDLLVYLLAHACALVLLPSPCAYQVCGSPFYQICA 180  
b 121 SYLPTNVTDALRGSGANGLLRLRRVDDVYVHLLARCALFVLVAFPCAYQVCGPPYQLCA 180  
Y 181 TTDIWPVSASVRYRTPVGRNFTNRLFQIKSGSSROEAPKPLALPSRGTKHLSTSTS 240  
b 181 ATQARPPPHAS-GPFRRLG-----CERAWNSHREAGVPLGLFAPGARRGGSASRS 231  
Y 241 VPSAKACVCPYRVEEGP-----HQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292  
b 232 LPLPKRPRGAAPPEPRTVQGGSHAPGRTGRGSDRGFCVVSFAR-----PAEATSLE 286  
Y 293 GKVSLSLS-GSVCKKHXPSTSLSPRONAFOLRP-FIETRHFLYSGDGOERLNPSF 350

Db 287 GALSSTRHSHSVGRQHHAGPPSTSRPPRWDTPCPVYAETKHFLYSSGD-KEQLRPSF 345  
QY 351 LLSNLQPNLTGARRLVEIFLGSRPRTSGPLCRTHRLSRRYQMRPLFOQLLVNHAECQY 410  
Db 346 LLSLSRPSLTGARRLVETIFLGSRPMPGTPRRLPRLPQRYQWQMRPLFLELLGNHACPY 405  
QY 411 VLLRSHCHCEFTANQQVTDAL-----NTSPHLMDLRLRLHSSPMQVY 452  
Db 406 GVLLKTHCPRLAA---VTPAAGVCAREKPGQSVAAPEEBDDPRLLVQLLRQHSSPMQVY 462  
QY 453 GFLRACLCKVYSASLWGTNRHNERFFKNLKKFISLGYKLSLQELMMKMKVEDCHWLRS 512  
Db 463 GFVRAELRELVPGLGWSRHNERRFLNTKKFISLGHAKLSLQELTWKMSVRDCAMLR 522  
QY 513 SPGRDVPAAEHLRBERILATFLWMDTYVQLRSEFFYFTESTFQKNLFFYRKSVMWS 572  
Db 523 SPGVGCVPAAEHLRBERILAKFLWMSVYVVELLRSEFFYFTESTFQKNLFFYRKSVMWS 582  
QY 573 KLSIGVROHLERVELRSLQSEVRRHODTWLAMPICRLRPIPKPGLRPIVNMYSMGT 632  
Db 583 KLSIGIRQHLKEVQLRELSEAEVQREARPAALLTSRLRPIPKPDGLRPIVNMVYVGA 642  
QY 633 RALGRKQKHQFTQRLKILFSLMNLRYTKHPLMGSSVGLGNDIYRTWRAFLVLRALDQ 692  
Db 643 RTFRREKRAERLTSRVKALFSLNRYERARRPCLLGASVGLDDIHRARWTFVLVRADQP 702  
QY 693 TRPMYFVKADVTGAYDAIPQGLKVEVWAMIRHSESTYCIQYAVVRRDSQGVHKSFRR 752  
Db 703 PELYFVKVDTGAYDTPQDELTEVIAIILK-PQNYICVRYAVVQKAAHGHVYKAFKS 761  
QY 753 QVTTLSDLQPYNGQFLKHLQSDASALRNSVVISQISMNNESSSLDFPFLHFLRHSVVK 812  
Db 762 HYSTLTDLPYMRQFVAHLQET--SPLRDVAVVIEQSSSLNEASSGLDFVFLRFMCHAVR 819  
QY 813 IGRDVCYTCQGGIPQGSLSLTLCSLFCFGMENKLPFAEVORDGLLRFVDDFLVTPHLQ 872  
Db 820 IRKSVYQCGGIPQGSLSLTLCSLCYGDMMENKLPFAGIRRDGLLRLVDDFLVTPHLTH 879  
QY 873 AKTFLSTLVHGVPEYGCMLNLOKTVYVFPFPGTLGGAAPYQLPAHCLFPWCGLLDTQT 932  
Db 880 AKTFLRTLVRGVPEYGCVMNLRKTVVNFVEDEALGGTAFVQMPAHGLFPWCGLLDTRT 939  
QY 933 LEVFCYSGYACTSIKTSITFSQSVKAGTWENKLLSVLRKCKHGLDLOVNSLQTVCI 992  
Db 940 LEVQSDYSYARTSRASLTFRNGFKAGNRMRKLVGVLRKCHSLFDLQVNSLQTVCT 999  
QY 993 NIYKIFELQAYRFHACVQLPFDQVRVKNLTPFLGIISSQASCCYAILKVNPGMTLKAAS 1052  
Db 1000 NIYKILLQAYRFHACVQLPFPHQVQWKNPTFFLAVISDTASLCVSLKAKNAGMSLGAK 1059  
QY 1053 GS---PPPEAAHMLCYQAFLLKLAHSHVYIKCLGLPRTAQKLLCRKLPKPEATMILKAAA 1109  
Db 1060 GAAGPLFSEAVQWLCHQAFLLKLTNRVTVYVPLLSGLRTAQTSRKLPGLTTTFALEAAA 1119  
QY 1110 DPALSTDFOTILD 1122  
Db 1120 NPALPSDFKTILD 1132

RESULT 10

AAB99930

ID AAB99930 standard; Protein; 1132 AA.

XX AAB99930;

DT 26-SEP-2001 (first entry)

XX Human telomerase protein sequence SEQ ID NO:31.

XX Differentiation; heart muscle cell; cytokine; transcription factor;

XX Proliferation; surface antigen; heart disease; cardiomyocyte;

KW Bone marrow; umbilical blood cell; heart muscle degeneration;

myocardial infarction.

Homo sapiens.

WO200148150-A1.

05-JUL-2001.

02-NOV-2000; 2000WO-JP07741.

28-DEC-1999; 99JP-0372826.

28-FEB-2000; 2000WO-JP01148.

(KYOW ) KYOWA HAKKO KOGYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

Yamada Y;

WPI; 2001-425655/45.

N-PSDB; AAH44366.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease

Claim 146; Page 137-141; 187pp; Japanese.

The present invention describes cells originating in bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. Also described are: (1) cardiomyocytes produced by the differentiation of the cells; (2) a method for carrying out the differentiation into cardiomyocytes, regulated by a promotional and/or inhibitory factor; (3) a method for the differentiation of the cells into cell types other than cardiomyocytes; (4) drug compositions promoting the formation of heart muscle and regeneration of heart tissue which contain the cells; (5) a method for the production of antibodies which recognise the cells, especially antibodies which recognise a surface antigen on the cells; (6) a method for screening factors which promote the proliferation of the cells; (7) a method for immortalising the cells by expressing telomerase in them; (8) drug compositions for the treatment of heart disease which contain the immortalised cells; and (9) cell-free supernatant from the culture of the cells and its use in promoting their differentiation into cardiomyocytes. The cells are used in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction and in the study of cardiomyocyte differentiation. AAH44351 to AAH44409 and AAB99915 to AAB99935 represent sequences used in the exemplification of the present invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MPRAPRCVAVSLRSRYREVWVPLATFVRRLGPRRLVOPGDPKIVRTLVAAQCLVCMHW 60

1 MPRAPRCVAVSLRSRYREVWVPLATFVRRLGPRRLVOPGDPKIVRTLVAAQCLVCMHW 60

61 GSQPPADLSHQVSSKELVARVQRLCERNERNVLAQFELINEARGGPPNMAFTSSVR 120

61 DARPPPAAPSFQVSCLELVARVQRLCERNERNVLAQFELINEARGGPPNMAFTSSVR 120

121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPFLVQICA 180

121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPFLVQICA 180

181 TTDDIWPSVASYRTPRVPGRNFTNLRFLQIKSSSRQEAAPKPLALPSGRTKRLHLSLTS 240

181 ATOARPPPHAS- GPRRLG- -----CERAWNHSVREAGVPLGLPAPGARRGGASRS 231

241 VPSAKARCYVPVRVEEGP- -----HRQVLPTPSGKSW- VPSPARSPFVPTAKDLSSK 292

232 LFLPKPRRGAPEPTEPVQGSWAHPGTRGFSDRGFCVSPAR- -----PAEEATSLE 286

QY 293 GKVSDDLSL- GSVCKKXKPSSTSLSPRONAFQURP- FIETRHFLYSKGDQGERINPSF 350  
 DB 287 GALSGRSHSPSVGRCHAGAPPSTSRPPRMDTPCPVYAEAKHFLYSSGD- KEQLRPSF 345  
 QY 351 LLSNLQNLGARRLVEIIFLGSRRPTSGPLCRTHLSRRYQWMPPLFOQLLVNHAEOY 410  
 DB 346 LLSLRPSLTGARRLVETIFLGSRRPMPGTPLRLPORYQWMPPLFLELLGNHAQCEY 405  
 QY 411 VRLRSHCRFRTANQQVTDAL- -----NTSPPHLMDLLRLHSSPWQY 452  
 DB 406 GVLLKTHCPRLAA- ---VTPAAGVCAREKPGQSVAAPEEDTDFRILVQLLRQHSPPWQY 462  
 QY 453 GFPRACLCVKVVSASLWGTNRERRFKNLKPTISLGKYGKLSLOELMMKMKVEDCHWLS 512  
 DB 463 GFVRACURRLVPPCLWGSRRNERRFURNTKKFISLGHAKLSLOELTWQSVRDCAWLR 522  
 QY 513 SPGKDRVPAAEHRLRERILATFLVLMDDTYVQLLRSFFIITESTFQKRLFFYKSVMS 572  
 DB 523 SPGVGCVPAAEHRLRERILAKFLHMLMSVVVVELLSFPYVTTETTFQKRLFFYKSVMS 582  
 QY 573 KLSIGVROHLRVRRLRISQBEVRHODTWLAMPICRLRFPKPNGLRPIVNMYSMGT 632  
 DB 583 KLSIGVROHLRVRRLRISQBEVRHODTWLAMPICRLRFPKPNGLRPIVNMYSMGT 642  
 QY 633 RALGRKQAQHFQRLKTLFSLMNLNERTKHPHLMGSSVLGMNDIVRTWRAFLVRALDQ 692  
 DB 643 RTPRERKGAERLTSRVKALFSLVNERARRPGLGASVLGLDDIHRANETFLVRAQDP 702  
 QY 693 TPRMYFKADVGTGAYDAIPGKLVVVVNMIRHSESTYCIROYAVVRDSDQGVHKSFR 752  
 DB 703 PPELYFYKVDVTGAYDTIPQDRLTEVIAIIR- PONTYCVREYAVVQKAAHGVKAFKS 761  
 QY 753 QVTLSLQPYMGQFLKHLQSDASALRNSVWIEQSI-SMNESSSLDFELHFLRHSVVK 812  
 DB 762 HVSLTDLQPYMGQFVAHLOET- --SPERDAVIEQSSSLNEASSGLFVFLRPMCHAVR 819  
 QY 813 IGDRCTYQCGIIPQSSSLTSLCSLFCGDMENKLFABVORDGLLIRFVDDFLLVTPHLQD 872  
 DB 820 IRGKSVYQCGIIPQSSSLTSLCSLFCGDMENKLFAGIRRDGLLIRLVDDFLLVTPHLTH 879  
 QY 873 AKTSLTLVHGVEYCCMINLQTVNFPVEPCTGGAPYOLPAHCLFPMWGLLDDTCT 932  
 DB 880 AKTFLRLTVRGVPEYGCVMNLRTVNFVVEDEAGGTAFVQMPAHGLFPWCGLLDDTCT 939  
 QY 933 LEVPCDYSYQAQTSIKTSTFTFQSVFKAGTMRNKLISVLRKCHGLFLLDQVNSLQTVCI 992  
 DB 940 LEVQSDYSYVARTSIRASLTFRNGFKAGRMNRKLFGLVLRKCHSLFLDLQVNSLQTVCT 999  
 QY 993 NIYKIFLQAYRFHACVILQFPDQVRKNTLTFPLGLISSQASCCVAILKVPKPGMTLKAS 1052  
 DB 1000 NIYKIFLQAYRFHACVILQFPDQVRKNTLTFPLGLISSQASCCVAILKVPKPGMTLKAS 1059  
 QY 1053 GS- --PPEEAHMLCYQAEFLKLAASHVLYKCLLGLRLTAQKLLCKRKLPEATMTILKAA 1109  
 DB 1060 GAAGPLSEAVOMLCHQAEFLKLTTRHRTVYVFLGSLRTAQQLSKLPGTTLTALEAA 1119  
 QY 1110 DPALSTDFQILD 1122  
 DB 1120 NPALPSDFKILD 1132

RESULT 11

AAG64329 standard; Protein; 1132 AA.

ID AAG64329

XX AAG64329;

AC AAG64329;

XX 24-SEP-2001 (first entry)

DT 24-SEP-2001 (first entry)

XX Human protein #2.

DE Human protein #2.

XX Angiogenesis; cardiant; cell differentiating agent; bone marrow;

KW Angiogenesis; cardiant; cell differentiating agent; bone marrow;

heart muscle cell; heart disease; human.

Homo sapiens.

WO200148149-A1.

05-JUL-2001.

28-FEB-2000; 2000WO-JP01148.

28-DEC-1999; 99JP-0372826.

(KYOW ) KYOWA HAKKO KOCYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K;

WPI; 2001-418252/44.

N-PSDB; AAH49601.

New adult bone marrow-originated cells capable of differentiating into heart muscle cells, applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration .

Disclosure; Pages 128-134; 158pp; Japanese.

The present invention relates to cells isolated from bone marrow, which are capable of at least differentiating into heart muscle cells. The cells are applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration. The present sequence was used to illustrate the present invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 1 MTRAPCPAVRSLRGRYREVLPATFVRLGPEGRRLVQPGDKIYTLVAQCLVCMHW 60  
 1 MPRAPCRVRSLLRSHYREVLPATFVRLGPGQWRLVQRGDPAAFRALVAQCLVCPW 60  
 61 GSOPPPADLSFHVSSIKELVARVQRLCERNVNLAFGPELLNEARGPPMAFTSSVR 120  
 61 DARPPPAAPSPFQVSCLELVARVQRLCERAKVNLAFGALLDARGGPEAFTSSVR 120  
 121 SYLFTVETIRVSGAMLLSRVGDLLVLLAHCALYLLVPPSCAYQVCGSFLYQICA 180  
 121 SYLFTVTDALRGSGAWGLLRRVGDVLLVHLLARCALPVLVAPSCAYQVCGPPLYLGA 180  
 181 TTDIWPVSASVYRTPRVGNFTNRLFLQIKSSSQEAPKPLALPSSGTRKHLSTST 240  
 181 ATQARPPPHAS-GRRRLG-----CERAMNHSVREAGVPLGPAPARRRGSASRS 231  
 241 VPSAKKARCPVPRVERGP-----HROVLPTPSGKSW-VPSAPSPVEVPTAEXDLSSK 292  
 232 LPLPKRPRRGAPEPEPTVPGQSWAHPCGTRGSPDRGFCVVSAPR----PAEATSL 286  
 293 GKVDLSLS-GSVCKKPKSTSLAPPQONAFQLRP-FIETRHFLYRGDQGERLNSP 350  
 287 GALSGRHSHPSVGRQHAGPSTSPRPWDTPCPVVAETKHFLYSSGD-KEQLRPSF 345  
 351 LLSNLQNLGTARRLVEIIFLGSRRPTSGPLCTHLSRYWQMRPLFOCLLVNHAECQ 410  
 346 LLSRLRSLGTARLVEIIFLGSRRPWPPTPRLPRLPORYWQMRPLFLELLGNAQCPY 405  
 411 VLLRSHCRFTANQOVTDAL-----NTSPPHMDLLRLHSSPWQVY 452  
 406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPWQVY 462  
 453 GFLRACLCVVSASLWGRNRRFFNLKFKSLGKYKLSLOELMWKVEDCHWLRS 512  
 463 GFVRACLRRLVPPGLWGRNRRFFNLTKFISLGHAKLSLOELTWKSVRDCAWLRR 522

QY 513 SPGKDRVPAEAHRLRERILATFLFPLMDTVVQLLSFFVITESTFQKRLPFYRSVMS 572  
 DB 523 SPGVCVPAEAHRLRERILAKFLHMLSVTVVLLRSFFVITETITFQKRLPFYRSVMS 582  
 QY 573 KLSIGVRQHLRLVRLRELQSEVRRHQDTWLAMPICRLFIKPNGLRPIVNMYSYMG 632  
 DB 583 KLSIGIRQHLKRVQLRSELSAEVRQREARPAALLTSRLRFIPKPDGLRPIVNDYVGA 642  
 QY 633 RALGRKQAOCHFTORLKTFLSMLNVERTKPHLWGSVLGMNDIYRTAFVLRVRLDQ 692  
 DB 643 RTFREKKAERLTSRVKALFVNLNERRRPGLLGASVLGDDIHRAWRTFVLRVRAQDP 702  
 QY 693 TPRMYFVKADVTGAYDAIPQGLVEVVANMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
 DB 703 PELYFVKVDVTGAYDTIPQRLVEVIAIIPONTYCVRRYAVVQKAHGHVRKAFKS 761  
 QY 753 QVTTLSDLOPYMGQFLKHLQDSASALRNSVWIEQSISSMESSSSLFDFELHRLHVVK 812  
 DB 762 HVSTLTDLQPYMROFVAHLOET--SPLRDVAVVIEQSSSLNEASSGLFDFELRMCHHAVR 819  
 QY 813 IGDRCYTCQGIPOGSSSLTLLCSLCFGDMENKLFABVORDGLLLRPVDDFLLVTPHLDQ 872  
 DB 820 IRGKSYVQCGIPQGSILSTLLCSLCYGDENKLFAGIRDRGLLLRLVDDFLLVTPHLLTH 879  
 QY 873 AKTFLSTLVHGVPGYGCINLQKTVNPFVPEPGTILGGAAPYQIPAHCLFPWCGLLDTQT 932  
 DB 880 AKTFLRTLVRGVPYGCYVNLKTVNPFVPEDEALGTAFAVQMPAHGLFPWCGLLDTT 939  
 QY 933 LEVFCDSYGAQTSIKTSLTFSQVFKAGKTMNKLLSVLRKLGHLFDLQVNSLOTVC 992  
 DB 940 LEVQSDYSYARTSIRASLTFRNFGFKAGRNWRKLFVLRKLGHLFDLQVNSLOTVCT 999  
 QY 993 NIYKIFLLQAYRFHACVQLPFDQVRNLTFFGIITSSOASCYAILKYNFGMTLKAS 1052  
 DB 1000 NIYKILLQAYRFHACVQLPFDQVWKNPTFFELRVISDTASLCYSILKAKNAGMSLGAK 1059  
 QY 1053 GS---FPPEAAHLCYQAFILKLAHSHVYIKLIGPLRTAKLCKLPEATTLKAAA 1109  
 DB 1060 GAAGPLPSEAVQWICHOAFLLKLTTHRVTVYVPLLSURTAQTQLSRKLPGTTTLALEAAA 1119  
 QY 1110 DPALSTDFQTILD 1122  
 DB 1120 NPALPSDFKTILD 1132  
 RESULT 12  
 AAG64859  
 ID AAG64859 standard; Protein; 1132 AA.  
 XX AAG64859;  
 XX AC AAG64859;  
 XX DT 21-SEP-2001 (first entry)  
 XX DE Heart muscle cell differentiation related protein SEQ ID NO: 31.  
 XX KW Heart muscle cell; human; cell differentiation; heart disease.  
 XX OS Homo sapiens.  
 XX PN WO200148151-A1.  
 XX PD 05-JUL-2001.  
 XX PF 27-DEC-2000; 2000WO-JP09323.  
 XX PR 28-DEC-1999; 99JP-0372826.  
 XX PR 28-FEB-2000; 2000WO-JP01148.  
 XX PR 02-NOV-2000; 2000WO-JP07741.  
 XX PA (KYOW ) KYOWA HAKKO KOCYO KK.  
 XX PI Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;  
 PI Yamada Y;

WPI; 2001-425656/45.  
N-PSDB; AAD468235.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease -

Claim 87; Page 143-147; 183pp; Japanese.

The present invention provides cells originating in the human bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. These cells are useful in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction, and the study of cardiomyocyte differentiation. The present sequence is a protein described in the exemplification of the invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
est Local Similarity 62.4%; Pred. No. 0;  
atches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVAVSLRSYREVMPLATVRLGEGRRLLVQPGDKIYRILVACCLVCMHW 60

1 MPRAPRCVAVSLRSYREVMPLATVRLGEGRRLLVQPGDKIYRILVACCLVCMHW 60

61 GSOPPPADLSFQVSSKELVARVQRLCERNERNVLAQFELLNEARGGPPMAFTSVR 120

61 DARPPPAAPSPQVSCVCLVAVQLQCBGAKVLAQFALLDQARGGPEAFTSVR 120

121 SYLNTVETLRVSGAWMLLSRVGDDLLVYLALCALVLLPSPCAYQVCGSPYQICA 180

121 SYLNTVTDALRGSGAWMLLSRVGDDLLVYLALCALVLLPSPCAYQVCGSPYQICA 180

191 TTDWPSVASYRTRPVGRNFTNLRFLQIKSSQERAPKPLALPSRGTKEHLSLTS 240

191 ATQARPPPHAS-GPFRRLG-----CESAMHNSVREAGVPLGAPARRRGSGASRS 231

241 VPSAKKARCYPPVPRVREGP-----HRQLVLPSPGKSM-VPSPARSPVPTAEKDLSSK 292

232 LPLKPRRGRGAPEPTPVQGSNAWHPGTRGSDRGFCVVSFAR-----PAEATSLE 286

293 GKVSGLSLG-SGVCKHKPSTLSLSPRONAQQLRP-FIETPHFLYSRGDQERLNPSF 350

287 GALSQTRHSHPSVGRQHHAGPPSTSPRPDPWTPCPBPVVAETRHFLYSRGD-KEQLRPSF 345

351 LLSNLQPNLTGARRLVETIFLGSRPSTSGPLCRTHLSRRYWOMRPLFOQLLVNHAECQY 410

346 LLSLRPSLTGARRLVETIFLGSRPWMPGTRPLRPLQRYWQMRPLFLELLONHAQCPY 405

411 VRLLRSHCRFTANQQVTDAL-----NTSPFHLMDLLRLHSSPWQVY 452

406 GVLLKTHCPLRAA--VTPAAGVCAKPKQGSVAAPPEEDTPRRLVQLLRHSSPWQVY 462

453 GFLRACILKVVASLWGRTRNRRFFKQLKFFSLGKYGLSLQELMWMKVEDCHNLS 512

463 GFVACILRLVPPGLWGRNRERFLNRTKFFSLGKHAKLSQELTWKSVRDCAWLRR 522

513 SPGRDVPAAEHRILREILATFLFWMMDTVVQLLRSFFYITESTFQKRLFFYKSVMS 572

523 SPGVGCVPAEHRILREILAKFLHLMSVTVVLLRSFFVYITETTFQKRLFFYKSVMS 582

573 KLOIGVROHLERVRLELSQEVRRHQDTWMLPICRLRFTPKPGLRPIVNMYSMTG 632

583 KLOIGIRQHLKRVQLRELSAEVRRQREARPAALLTSRLRFTPKPDGLRPIVNMVYVGA 642

633 RALGRKQAOHFTORLTLFSLNLYERTKPHLMGSSVLGMNDIYETWRAFLVRLALDQ 692

643 RTFREKRAELTSRVKALSVNLNRYEARPGVLGASVLGDDIHAWTFLVRLAQDP 702

693 TPRMYFVKADVGTAYDAIPQOKLVEVVANMIRHSESTYCIQVAVVRDSSQGVHKSFR 752

Db 703 PPELYFKVDVTGAYDTIPQDRLTEVIAIILK-PONTYCVRRYAVVQKAAHGHVKAPKS 761  
Qy 753 QVTTLSDLOPYMGQFLKHQDSADALRNSVIEQISIMNESSSLDFDFLFLHSHSVK 812  
Db 762 HVSTLTDLOPYMRFQVAHLQET--SPLRDVAVIEQSSSLNEASSGLFDVFLRFMCHHAVR 819  
Qy 813 IGDRCTYQCGIGIQSGSSLTLLCSLCFCGDMENKLFABVQRDGLLIRFVDDFLVTPHLDQ 872  
Db 820 IRKSYVQCGIGIQSGSSLTLLCSLCFCGDMENKLFABVQRDGLLIRFVDDFLVTPHLDQ 879  
Qy 873 AKTFLTLVHGVPEYVGCMLNOKTVNPPVEPTLGGAAPIQLPAHCLFPMCGLLDQT 932  
Db 880 AKTFLTLVHGVPEYVGCMLNOKTVNPPVEPTLGGAAPIQLPAHCLFPMCGLLDQT 939  
Qy 933 LEVPCYSGVAOTSISITFQSVFKAGTKMENKLLSVLRKLCGSLFLDLQVNSLQTVCI 992  
Db 940 LEVQSYSSVARTSIRASLTFRGFKAGNMRKLVGVLRKCHSLFLDLQVNSLQTVCI 999  
Qy 993 NIYKIFLQAYRFHACVIOQPFQDQVRKNTLFPFLGIISSQASCCYAILKVRNPGMTLKAS 1052  
Db 1000 NIYKILLQAYRFHACVIOQPFQDQVRKNTLFPFLGIISSQASCCYAILKVRNPGMTLKAS 1059  
Qy 1053 GS---FPPEAAHWLVCQAEELKLAHNSVYKCLIGFLTAQKLCRKLPEATMTILKAAA 1109  
Db 1060 GAAGPLFSEAVOMLCHQAFLLKTRHRTVYVPLLSGLRTAQTSRKLPGTTLTALAAA 1119  
Qy 1110 DPALSTDFQITILD 1122  
Db 1120 NPALPSDFKTLILD 1132

RESULT 13  
AAE29226

ID AAE29226 standard; Protein; 1132 AA.

XX AAE29226;

XX AC

XX DT 27-JAN-2003 (first entry)

XX DE Human telomerase reverse transcriptase (TERT).

XX KW Carbohydrate antigen; alpha(1,3)galactosyltransferase; alpha1,3GT; TERT;

XX KW transgenic; alpha(1,2)fucosyltransferase; alpha1,2FT; human; enzyme;

XX OS Homo sapiens.

XX WO200274948-A2.

XX PD 26-SEP-2002.

XX PF 21-MAR-2002; 2002WO-CA00378.

XX PR 21-MAR-2001; 2001US-277811P.

XX PA (GERO-) GERON CORP.

XX PI Denning C, Clark AJ, Schiiff JN;

XX DR WPI; 2002-759895/82.

XX DR N-PSDB; AAD46821.

XX PT Mammalian cells, useful for producing animal tissues with carbohydrate

XX PT antigens that are compatible for transplantation into human patients -

XX PS Disclosure; Page 34; 71pp; English.

XX CC The invention relates to animal tissues with carbohydrate antigens that

XX CC are compatible for transplantation into human patients. The mammalian

XX CC cell is inactivated homozygously for expression of alpha(1,3)galactosyl-

XX CC transferase (alpha1,3GT) gene and comprises a transgene for alpha(1,2)-

XX CC fucosyltransferase (alpha1,2FT). It is useful for producing animal tissue

XX CC with carbohydrate antigens that are compatible for transplantation into



human patients. The present sequence is human telomerase reverse transcriptase (TERT) used in the invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 23; Length 1132;  
Identical Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
1 MTRAPCRVAVSLRSRYREVMPLATFVRRLGPEGRRLVQPGPKIYRTLVAAQCLVCMHW 60  
1 MTRAPCRVAVSLRSRYREVMPLATFVRRLGPEGRRLVQPGPKIYRTLVAAQCLVCMHW 60  
61 GSQPPADLSHVSLSKELVARVYVORLCERNENVLAFPELLNARGGPPMAFTSVR 120  
61 DARPPAAPSPFRQVSCLELVARVQRCERGAKNVLAFAFALLDARGGPPPAFTSVR 120  
121 SYLPTNVTIETLRVSGAMMLLSRVGDDLLVYLLAHALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPTNVTIETLRVSGAMMLLSRVGDDLLVYLLAHALYLLVPPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASVYRTPRVGRNFTNLRFLQOKSSSQEAPKELALPSRGTKYRHLSTSTS 240  
181 ATQARPPPHAS--GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRGSSASRS 231  
241 VPSAKKACYPVPRVEEGP-----HRQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292  
232 LPLPRPRGAAPFERFVQGSWAHPGRFGRGFCVVSFAR-----PAEATSLE 286  
293 GKVDLSLS-GSVCKKHPSTSLSPRQNAFOLRP-FIETHFLYSRGDQGERLNPSF 350  
287 GALSCTRSHSPSVGRQHAGPSTSRPRPMDTFCPPVYAEKTHFLYSSGD-KEQLRPSF 345  
351 LLSNLQPNLTGARRIVEIIFLGSRRPTSGPLCTRHSRYWOMRPLFOQLLVNHAECV 410  
346 LLSLRPSLTGARRIVEIIFLGSRRPMTGPRPLPQRYWQMRPLFELLGHNHACPY 405  
411 VRLRSHCRFRFANOQVTDAL-----NTSPHLMDLRLHSSPWQVY 452  
406 GYLLKTHCFLRAA---VTPAGVCAREPQGSVAAPBEEDTPRRLVQLLRQSSPWQVY 462  
453 GTRACLCKVVSASLWGRHNEREFKXKFI SLGKYGKLSLOELMKMKWCKDCHWLS 512  
463 GVRACLRRLVPPGLWGRHNEREFKXKFI SLGKHAKLSLOELTWKMSVRCDAWLR 522  
513 SPKORVFAAEHRLRERILATFLFWMYVQLIRGFFYITESTFQKNRLLFYRKSWS 572  
523 SFGVGCVPAAEHRLRERILAKFLHLMVYVVELLRGFFYITESTFQKNRLLFYRKSWS 582  
573 KLOSTGVROHLERVLRLBSQEVREHDOOTWLAMPI CLRIPKPNGLRPIVNNYSMTG 632  
583 KLOSTGIROHLKRVQLRELSEAEVQREAPALLTSRLRFPKPDGLRPIVNNYVVG 642  
633 RALGRKQAOHTQRLKTLFLSMYNTERTKPHLMGSSVLGMNDIYRTWRFAVLRLVALDQ 692  
643 RFRREKRAERLTSAVKALFSLVNYERARRPCLLGASVLGDDTHRAWRTFVLEVRADP 702  
693 TRPMYFVADVTGAYDAIPQGLKLVENVANIRHSESTYCIQYAVVRDQGGQVHKSFRR 752  
703 PPELYFVKVDVTGAYDTIPQDLTEVIAIHK-PONTYCVRRYAVVQAAHGHVYKAFKS 761  
753 QVTLSDLOPYMGQFQKHLQSDASALANSVVEIOSISMNESSSLDFFLHFLRHSVVK 812  
762 HVSITLTLQPYWRQFAHLQET--SPLDAAVVEIOSSSLNEASSGLFDVFLRNFCHAVR 819  
813 IGDRCYTCQGI:PGQSSSLTLCSCFCGDMENKLPFAEVRDGLLRFDVFLVTPHLQD 872  
820 IRGKSVYQCGIG:PGQSSSLTLCSCFCGDMENKLPFAEVRDGLLRFDVFLVTPHLTH 879  
873 AKTFLSTLVHGVPEVGCMLNOKTVNPFVPGTLGGAAPQLPAHCLFPWCGLLDTQT 932  
880 AKTFLTLVRGPEVGCVMNLRKTVNPFVDEALGGTAFVQMPAHGLFPWCGLLDTRT 939

Qy 933 LEVFCDSGYAOTSIKTSLTFSQVFKAGTWENKLLSVLRKCHGLFDLQVNSLQTVCI 992  
Db 940 LEVQSDYSYARTSIRASLTFNRGPAAGNWRKULFGVLRKCHSLFDLQVNSLQTVCT 999  
Qy 993 NIYKIFLQAYRFHACVQLPFDQVRVKNLTFFLGIISQASCCVAILKVRKPGWTLKAS 1052  
Db 1000 NIYKILLQAYRFHACVQLPFDQVRVKNLTFFLGIISQASCCVAILKVRKPGWTLKAS 1059  
Qy 1053 GS---PPPEAAHWLQYQAEFLKLAHSAVYKCLLGLPLTAOKLLCRKLPEATMTILKAAA 1109  
Db 1060 GAGCPUPSEAVQWCHQAEFLKLTTRHRTVYVFLGLSLRTAQTQLSRKLPGTTLTLEAAA 1119  
Qy 1110 DPALSTDFOTILD 1122  
Db 1120 NPALPSDFKTILD 1132  
RESULT 14  
AAU72735  
ID AAU72735 standard; Protein; 1132 AA.  
XX AC AAU72735;  
XX DT 09-APR-2002 (first entry)  
XX Human telomerase reverse transcriptase (TERT).  
XX Telomerase reverse transcriptase; TERT; cytosolic; apoptosis;  
XX cell growth inhibitor; antisense oligonucleotide;  
XX antisense technology.  
XX Homo sapiens.  
XX WO2001:88198-A1.  
XX 22-NOV-2001.  
XX 15-MAY-2001; 2001WO-US15774.  
XX 16-MAY-2000; 2000US-0572423.  
XX 07-DEC-2000; 2000US-0733294.  
XX (ISIS-) ISIS PHARM INC.  
XX Monia BP, Gaarde WA, Freier SM, Wanciewicz E;  
XX WPI; 2002-075321/10.  
XX N-PSDB; AAS96607.  
XX New compound targeted to nucleic acid molecule encoding telomerase  
XX transcriptase (TERT), which specifically hybridises with and inhibits  
XX expression of TERT, useful for modulating apoptosis and inhibiting cell  
XX growth -  
XX Disclosure; Page 100-105; 154pp; English.  
XX The invention describes a compound, 8-50 nucleobases in length targeted  
XX to a nucleic acid molecule encoding human TERT (telomerase reverse  
XX transcriptase), where the compound specifically hybridises with and  
XX inhibits the expression of TERT. A series of oligonucleotides were  
XX designed to target different regions of the human TERT RNA. These were  
XX 20 nucleotides in length and composed of a central gap region consisting  
XX of ten 2'-deoxynucleotides, flanked on both sides (5' and 3' directions)  
XX by five-nucleotide wings. The wings were composed of 2'-methoxyethyl  
XX (2'-MOE) nucleotides. The compounds were analysed for their effect on  
XX human TERT mRNA levels by reverse transcriptase (RT)-polymerase chain  
XX reaction (PCR). The compound is useful for inhibiting the expression of  
XX TERT in cells or tissues, for treating a human having disease or  
XX condition associated with TERT, for modulating apoptosis, for inhibiting  
XX cell growth (preferably, cancer cell growth), in antisense therapy and  
XX for diagnostics and therapeutics. This is the amino acid sequence of  
XX human telomerase reverse transcriptase (TERT), described in the method  
XX of the invention.



nose, oesophagus, stomach, or intestine. The present sequence represents human TERT (Htert), which is given in the exemplification of the present invention. Htert is located to chromosome 5.

Sequence 1132 AA;

```
Query Match          59.4%; Score 3505; DB 24; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1  MTRAPRCFAVSLRSRVREVMPLATFVRLLGPEGRLVQPGPKIVRTIVAQCLVCMHW 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1  MPRAPRCFAVSLRSRVREVMPLATFVRLLGPEGRLVQPGPKIVRTIVAQCLVCMHW 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61  GSQPPADLSFHQVSSKELVARVVQBLCEENRVLAFLGELLNARGPPMAFTSSVR 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61  DARPPPAAPSFRQVSCLELVARVLQRCERGAKNVLAFLGELLNARGPPMAFTSSVR 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

121 SYLPTNTVETLRVSGAMLLSRVGDLLVYLLAHCALVLLVPPSCAYOVCGSPLYQICA 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SYLPTNTVDALRGSGWGLLRRVGGDDVHLLARCALFVIVAPSCAYQVCGPPLYQLGA 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

181 TTDIWPVSASYRTPRVGRNFTNLRFQQIKSSRQEPAPKPLALPSRGTKRHLSTSTS 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 ATQARPPPHAS-GPERRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSASRS 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

241 VPSAKKACYPVPRVEEGP-----HQVLPTPSGKSW-VPSPARSPEVPTAEKDLSSK 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
232 LPLPKRPRGGAAPERTPVQGGSWAHGPRTRGSPDRGFCVGPAP-----PAEATSLSE 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

293 GKVDLSLS-GSVCKHKPSSLSLPPRQNAFQLRP-FIETRHFLYSRGDGOERLNPSF 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
287 GALSCTRSHSPSVGRQHAGPPSTSRPPMDTPCPVVAETKHFVYSSGD-KEQLRPPSF 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

351 LLSNLQPNLTGARLVEIFIGSRPRISGPRCTRHLSRBYQWQWELFQOLLVNAECQY 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 LSSLRPSLTGARRLVETIFIGSRPWPMPGTFRRLPRLPQRYQWQWELFLELLGNHAQCPY 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

411 VRLLRSHCRFRFANQQVTDAL-----NTSPHMLDLRLHSSPWQVY 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 GVLAKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSWPQVY 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

453 GFRLACLCKVVSASLWGRTHNERPFQNLKFIISLGYKGLSLQELMWKWKVDDCHWLRS 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 GFVRACLRLVPPGLWGRSHNERFLRNTKFIISLGHAKLSLQELTWKQSVRDCAWLRR 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

513 SPQKRVPAAEHRLRERILATFLWMDTVYVQLLRGFFYITESTFQKRLFFYRKSQVMS 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
523 SPQVGCVPAAEHRLRERILAKELHLMVSYVVELLRGFFYITESTFQKRLFFYRKSQVMS 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

573 KLSIGVROHLRLVRLSLSQEEVVRHQDTWLAMPICLRPIPKNGRLRPIVMSYSMTG 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
583 KLSIGIRQHLKRVQRELSAEVVRHQREARFALITSRLRPIPKPDGLRPIVMDYVYVGA 642
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

633 RALGRKQAEHFTQRLKTLFSLNRYERTKHPHLMGSSVLGMNDIYRTWRAFLVRLALDQ 692
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
643 RTFRREKRAERTSRVKALFSVINTERARRPGLLGASVLGLDDIHRWRTFVLVRAQDP 702
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

693 TPRIYFVKADVTGAYDAIPQGLVEVWVANNIRHSESTYCIQYAVVRDSDQGVHKSFR 752
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIHK-PONTYCVRRYAVVQKAAHGHVRKAFKS 761
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

753 QVTLSDLPYMGQFLKHLQDSALNSVVIQOSISMNESSSLPDPFLHLRHSVVK 812
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 HYSTLTDLQFYMRQFVAHLQET--SFLRDVAVVIEOSSLSNEASGLFDVFLRFMCHHAVR 819
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

813 IGDRCYTQCGIPQSGSLSTLCSLCFGDMENKLPFAEVRQDGLLIRFVDDFLVTPHLQ 872
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
820 IRGKSVYVQCGIPQSGSLSTLCSLCYGDWENKLPAGIRRDGLLRLVDDFLVTPHLTH 879
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

873 AKTFLSTLVHGVPEYGCMLNQKTVNPFVEPGTLGGAAPYQLPAHCLFPWCGLLDITQT 932
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
880 AKTFLRLVRGVPEYGCVVNLRKTVNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTRT 939
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: February 4, 2004, 14:10:19

Job time : 51 secs

```
Qy 933 LEVFCYSGYAQTSIKTSITFQSVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCI 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 LEVQSDYSSYARTSIRASLTFRGFKAGNMRKLFGLRLKCHSLFLDLQVNSLQTVCT 999
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 993 NIYKIFLQAYRFHACVIOLPFDQVRKXLTFFLGIISQASCCYAILKXNPGMTLKAS 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 NIYKILLQAYRFHACVIOLPFHQQVWKNPFFFLAVISDTASLCYSILKAKNAGSLGAK 1059
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1053 GS---PPPEAAHWLCYQAFLLKAAHSVIYKCLLGLPLTAQKLLCRKLPBATMTILKAAA 1109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 GAAGPLPSEAVQWLCHQAFLLKLTIRHRTVYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1110 DPALSTDFQTILD 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 NPALPSDFKTILD 1132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 4, 2004, 14:09:14 ; Search time 21 Seconds  
(without alignments)  
2260.609 Million cell updates/sec

le: US-09-042-460-2

ect score: 5901

ence: 1 MTRAPRCPAVRLLRSRYRE.....TILKAADPALSTDFQITLD 1122

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 328717 seqs, 423-0858 residues

al number of hits satisfying chosen parameters: 328717

imum DB seq length: 0

imum DB seq length: 2000000000

:-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : Issued Patents:AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep:\*
- 2: /cgn2\_6/prodata/1/iaa/5S-COMB.pep:\*
- 3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep:\*
- 4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep:\*
- 5: /cgn2\_6/prodata/1/iaa/6CTUS-COMB.pep:\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

alt No.	Score	Query Match	Length	ID	Description
1	3505	59.4	1132	3	US-08-851-843A-225
2	3505	59.4	1132	3	US-08-974-549A-2
3	3505	59.4	1132	3	US-08-854-050-225
4	3505	59.4	1132	4	US-09-430-323-101
5	3505	59.4	1132	4	US-09-128-354-2
6	3505	59.4	1132	4	US-09-575-321-2
7	3505	59.4	1132	4	US-09-052-919-2
8	3505	59.4	1132	4	US-08-912-951-2
9	3505	59.4	1154	3	US-08-974-549A-611
10	3505	59.4	1154	4	US-08-912-951-323
11	3505	59.4	1189	3	US-08-974-549A-613
12	3505	59.4	1189	4	US-08-912-951-325
13	3505	59.4	1200	3	US-08-974-549A-612
14	3505	59.4	1200	4	US-08-912-951-324
15	3505	59.4	1285	3	US-08-974-549A-600
16	3505	59.4	1285	4	US-08-912-951-314
17	3496	59.2	1132	3	US-08-974-549A-344
18	3496	59.2	1407	3	US-08-974-549A-628
19	3496	59.2	1407	4	US-08-912-951-334
20	2492.5	42.2	1003	3	US-08-851-843A-217
21	2492.5	42.2	1003	3	US-08-974-549A-336
22	2492.5	42.2	1003	3	US-08-854-050-217
23	2492.5	42.2	1003	4	US-09-430-323-217
24	2195	37.2	807	3	US-08-974-549A-5
25	2195	37.2	807	4	US-08-912-951-5
26	1686.5	28.6	364	4	US-09-417-485D-41
27	1506.5	25.5	564	3	US-08-851-843A-101

Sequence 267, App  
Sequence 101, App  
Sequence 101, App  
Sequence 40, Appl  
Sequence 602, App  
Sequence 316, App  
Sequence 603, App  
Sequence 317, App  
Sequence 604, App  
Sequence 318, App  
Sequence 605, App  
Sequence 319, App  
Sequence 10, Appl  
Sequence 606, App  
Sequence 320, App  
Sequence 47, Appl  
Sequence 202, App

ALIGNMENTS

RESULT 1

US-08-851-843A-225  
; Sequence 225, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6093809e1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/851,843A  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 225:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

08-851-843A-225

very Match 59.4%; Score 3505; DB 3; Length 1132;

est Local Similarity 62.4%; Pred. No. 0;

atches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCFAVASLRSRYREVMPLATEVRLGPEGRLVQGPDKIYRILVAQCVCVHW 60

1 MERAPRCFAVSLRSHYREVLPLATFVRLGPGQWELVQSGDPAFRAVLAQCVCVWP 60

61 GSQPPADLSHQVSSIKELVARVQBLCEENRVNLAQFELLNEARGPPWATTSVR 120

61 DARPPAAPSPFQVSCLEKELVARVQLQCEGAKNVLAQFALLDQARGPPEAFTSVR 120

121 SYLPTVITELRVSGAWMLLSRVGDDLLVYLLAHCALVLLVPPSCAYQVCGSPLYQICA 180

121 SYLPTVTDALRGSGAWMLLSRVGDDVYVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180

181 TTDIWPVSASYPTRPVGRNFTNLRFLOQIKSSROEAPKPLALPSRGTKHLSLTSTS 240

181 ATQARPPPHAS-GPERRLG-----CERAWNSVREAGVPLGLPAPGARRRGSSARS 231

241 VPSAKKACYPVPRVEGCP-----HQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292

232 LPLPKRPRGAAPPERTPVQGGSNWHPGTRGSPDRGFCVVSFAR-----PAEATLSLE 286

293 GKVDLSLS-GSVCKKHPKSPSTLSLPPRQNAFQLRP-FIETRHFLYSRGDQOERLNPSF 350

287 GALSCTRSHPSVGRQHEAGPSPSTRPPRMDTPCPVYAEATKHFLYSSGD-KEQLRPSF 345

351 LLSNLQPNLTGARLVELIFLGSPPRSGPLCRTHLSRYWQWRLPQOLLVNHAEQCY 410

346 LLSSLRPLTGARLVELIFLGSPPRSGPLCRTHLSRYWQWRLPQOLLVNHAEQCY 405

411 VLLASHCRFRFTANQQVTDAL-----NTSPPHLMDLLRLHSSPWQVY 452

406 GVLLKTHCPLRAA---VTPAAGVCAKEKPGQSVAAPEEDTPRRLVQLLRQSSPWQVY 462

453 GFLRACLCKVNSASLWGRNHRERFFKNLKKFISLKGKGLSKLSQELMWKVKYDDCHWLS 512

463 GFVRACLRLVPPGLWGRNHRERFLNKKFISLGHAKLSQELTWKNSVEDCAWLR 522

513 SPCKDRVPAASHRLRERILATFLWLMDDTYVQLLRFFVITESTFQKRLFFYKSVMS 572

523 SPGVGCVPAASHRLREELAKFLHLSVYVVELLRFFVITETTKQNLFFYKSVMS 582

573 KLSIGVGRQHLEVRVRLRELSQEVVRRHODTWLAMPICRLRFPKPNGLRPIVNSYSMTG 632

583 KLSIGIRQHLKRVQLRELSAEVQRHREARFALLTSRLRFIPKPDGLRPIVNMDDYVGA 642

633 RALGRKQAOHTQRLKTLFSLMAYERTKPHLMGSSVLGMNDIYRTWRAFLVRVRLDQ 692

643 RTFREKRAELTSRVKALFSLVNYERARRPCLLGASVLGDDIHRARWTFVLVRVRAQDP 702

693 TPRMYFVADVTGAYDAIPQGLVEVAVNMIHRSESTYICIRQAVVRRDSQGVHKSFR 752

703 PPELYFVKVDVTGAYDTT-PQDLREVIASIIK-PQNTVCVRRVAVVQKAAHGVKAFKS 761

753 QVTTLSDLQPYMGQFLKHLQSDASARNSVVEQSI SMNRESSLDFDLHFLRHSVVK 812

762 HVSTLTDLQPYNRQFVAHQET--SPLSDAVIEQSSSLNEASSGLFDVFLRMCHAVR 819

813 IGDRCYTCQCGIPQGSSTLCSLCFQDMENKLFAYEQRDGLLLRVDVDFLLVTPHLDQ 872

820 IRGKSVQCGIPQGSSTLCSLCFQDMENKLFAYEQRDGLLLRVDVDFLLVTPHLDQ 879

873 AKTFSLTLVHGVPYEGCMINQKTVNPNFVPGTLGGAAFYQLPAHCLFPWCGLLLDTQT 932

Db 880 AKTFELTVRGVPEYCCVNLKRTVVPFVEDEALGGTAFVQMPAHGLFPMCGLLLDTRT 939

Qy 933 LEVFCYSGYAOITSIKTSLTFSVFKAGTKWRNKLKSLVLRKCHGLFLDLOVNSLOTVCI 992

Db 940 LEVQSDYSYVARTSIRASLTFRNGFKAGNRMRKLFGLRLKCHSLFLDLOVNSLOTVCT 999

Qy 993 NIYKIFLLQAYRFHACVQLPFDORVRKNLTFFLGISSQASCCVAILKVNKPNWTLKAS 1052

Db 1000 NIYKILLQAYRFHACVQLPFPHQQWKNPFFFLVISTASLTCYSILKAKNAGMSLGAK 1059

Qy 1053 GS---PPPEAAHWLCYQAFLLKLAHSHSVYIKLGLPLTAQKLLCRKLPEATWTLKAAA 1109

Db 1060 GAAGPLPSEAVQWLCHQAFLLKLTNRHRTVYVPLGLSLRTAQQLSRKLPGLTTLALEAAA 1119

Qy 1110 DPALSTDFOTILD 1122

Db 1120 NPALPDSDFOTILD 1132

## RESULT 2

US-08-974-549A-2

; Sequence 2, Application US/08974549A

; Patent No. 6186178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin B.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/974,549A

; FILING DATE: 19-NOV-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/915,503

; FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US97/17618  
 FILING DATE: 01-OCT-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US97/17885  
 FILING DATE: 01-OCT-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph Ted  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002610US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO. 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 -08-974-549A-2

Query Match 59.4%; Score 3505; DB 3; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCRAVRLLRSRYEVNPLATFVRRLLGPEGRRLVQPDGPKIVRTVAOCLVCMHW 60  
 1 MTRAPRCRAVRLLRSRYEVNPLATFVRRLLGPEGRRLVQPDGPKIVRTVAOCLVCMHW 60  
 61 GSQPPADLSHFQVSSKELVARVQRI-CERNERNVLAFFGELLNEARGPPMAFTSSVR 120  
 61 DARPAPAPSPFQVSCLELVARVLQRCERGAKNVLAFFGELLNEARGPPMAFTSSVR 120  
 121 SYLNPVTIELRVSGAWMLLSRVGDDLLVLLAHALYLLVPSCAVQVCGSPLOYICA 180  
 121 SYLNPVTIELRVSGAWMLLSRVGDDLLVLLAHALYLLVPSCAVQVCGSPLOYICA 180  
 181 TTDIWPVSASVYRTPVGRNFTNLRFQIKSSSRQAPKPLPALPSGTRKRLSLTSTS 240  
 181 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAPGARRRGGASRS 231  
 241 VPSAKKACVYPRVVEEP-----HQVLPSPGKSW-VPSAPSPVEPTAEKOLSSK 292  
 232 LPLKRRRGAAPERTFVGGSWAHGPRTRGSDGFCVVSAR-----PAEATSLSE 286  
 293 GKVSLSLS-GSVCCCKHPSSTLSLPPRQNAFQLRP-FIETRHLYSRGDSGQERLNPSF 350  
 287 GALSCTRSHSPVSGHQHAGPPSTSRPRPNDTPCPVYATKHFLYSSGD-KEQLRPSF 345  
 351 LLSNLOPNUTGARRLVEIFLGSRRPTSGPLCRHRSRRYQWRPLFOQLLVNHAECQY 410  
 346 LLSLRPSITGARRLVEIFLGSRRPTSGPLCRHRSRRYQWRPLFOQLLVNHAECQY 405  
 411 VLLRSHCRFRNTANQVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
 406 GVLLKTHCPLRAA---VTPAGVCAREKYPQSVAAPEEDTPRRVLQLRQSSPWQVY 462  
 453 GFLRACLKCVSASLWGRTHNRRERFKNLKFISLGKYKLSLQELMWKMKVEDCHWLS 512  
 463 GFVRACLRLVPPGLWGRTHNRRERFKNLKFISLGKYKLSLQELMWKMKVEDCHWLS 522  
 513 SPQKORVPAERLBERILATFLWMTYVOLLRSFFYITESTFQNRLLFFYKRSVWS 572  
 523 SPQKORVPAERLBERILATFLWMTYVOLLRSFFYITESTFQNRLLFFYKRSVWS 582  
 573 KLSQIGVROHLRRLRSLRQSEVEHHDDTWLAMPICRLRPIKPNGLRPIVNNYSMGT 632  
 583 KLSQIGVROHLRRLRSLRQSEVEHHDDTWLAMPICRLRPIKPNGLRPIVNNYSMGT 642  
 633 RALGRKQKQCHTQRLKTLFSLMNVKTRKHPHLMGSSVIGNDIYRTWRFAVLRLADQ 692  
 643 RTRRKRABRLTSRVKALFSLVNYERARRRPLGLIGASVLGLDD:HRARWTFVLVRQAQP 702

RESULT 3

US-08-854-050-225  
 ; Sequence 225, Application US/08854050  
 ; Patent No. 6261836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cech, Thomas R.  
 ; APPLICANT: Lingner, Joachim  
 ; APPLICANT: Nakamura, Toru  
 ; APPLICANT: Chapman, Karen B.  
 ; APPLICANT: Morin, Gregg B.  
 ; APPLICANT: Hartley, Calvin  
 ; APPLICANT: Andrews, William H.  
 ; TITLE OF INVENTION: No. 6261836el Telomerase  
 ; NUMBER OF SEQUENCES: 225  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States of America  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/854,050  
 ; FILING DATE: 09-MAY-1997  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/851,843  
 ; FILING DATE: 06-MAY-1997  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/846,017  
 ; FILING DATE: 25-APR-1997  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/844,419

QY 693 TPRMYFVADVTGAYDAIPOQKLVVVVANNIRSESTYCIQYAVVVRDSQGVHKSFR 752  
 DB 703 PPELYFVKVDVTGAYDTIPQDLRTVIASTIK-PONTYCVRRYAVVQKAHGHVRKAFKS 761  
 QY 753 QVTLTSLDQPYMGOFKHLQSDASALRNSVWIEQSTSMNESSSLFDFPLHFLRHSVVK 812  
 DB 762 HVSTLTDLQPYMRQFVAHQET--SPLRDAVWIEQSSSLNEASSGLFDVFLRPMCHAVR 819  
 QY 813 IGRCYTOCQIGIPOGSSSLTSLCSLFCGDMENKLFVAVQSDGLLLRFDVDFLLVTPHLDO 872  
 DB 820 IRGKSYVQCQIGIPOGSSSLTSLCSLFCGDMENKLFVAVQSDGLLLRFDVDFLLVTPHLTH 879  
 QY 873 AKTELSTLHVGVPEYGCWMLNOKTVNFPVPGTLGGAAPYQLPAHCLPWCGLLLDTOT 932  
 DB 880 AKTELTLRVGVPEYGCWMLNOKTVNFPVPGTLGGAAPYQLPAHCLPWCGLLLDTOT 939  
 QY 933 LEVFCDSYGYAQTSLTFSQVFKAGKTMRLKLSVLEKCHGLFLDLQVNSLOTVCI 992  
 DB 940 LEVQSDYSSYARTSIRASLTFRNGFKAGRNRRKLFGLVLEKCHGLFLDLQVNSLOTVCT 999  
 QY 993 NIYKIFLLOAYRPHACVIOQLPFDQVRKNTLTPFLGIISQASCCYAILKVKNPMTLKAS 1052  
 DB 1000 NIYKIFLLOAYRPHACVIOQLPFDQVRKNTLTPFLGIISQASCCYAILKVKNPMTLKAS 1059  
 QY 1053 GS---FPEEAHMLCYQAFLLKLAHSAVYKCLLGLPLRTAQKLLCEKLPATMTILKAAA 1109  
 DB 1060 GAAGPLPSEAQVWLCHOAFLLKTRHRTVYVPLLSLRTAQTOLSKLPGTTLTLEAAA 1119  
 QY 1110 DPALSTDFQTILD 1122  
 DB 1120 NPALPSDFKTILD 1132

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00293005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 225:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

i-08-854-050-225

Query Match

Best Local Similarity 59.4%; Score 3505; DB 3; Length 1132;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAYRSLRSHRYEWMPLATFVRRLGPEGRRLVQPDKXIRTLVAQCCLVCMHW 60

1 MPAPRCRAVRSLLRSHRYEVLPLATFVRRLGQGRVLRQGRDPAAPRALVAQCCLVCPW 60

61 GSQPPADLSFHQVSSIKELVARVQRCLERNRYLAFGFELLNEARGPPMAFTSVR 120

61 DAREPPAAPSRFQVSCIKELVARVQRCLERNRYLAFGFELLNEARGPPMAFTSVR 120

121 SYLPTVETLVSGAWMLLSVGDLLVYLAHALVLLVPBSCAVCGSPLYQICA 180

121 SYLPTVTDALRSGAGLRLRVGDVLLVHLLARCALFVLVAPSCAYCVCQPPYUQIGA 180

181 TTDIWPVSASYPTRPVGRNFTNLRFLOQIKSSSRQEAQKPLALPSRGTGRHLSLTSS 240

181 ATQARPPPHAS-GPRRELG-----CERAWNHSVREAGVPLGLPAPGARRRGSASRS 231

241 VPSAKKARCYPVRVEGP-----HRQVLPTPSGKSW-VSPARSPEVPTAEKDLSSK 292

232 LPLPKRPRGAPEPERTPVGQSWAHQRTGRPSDRGFCVVSAPAR-----PAEATSLE 286

293 GKVSDLSLS-GSVCKHKPSTSLGPPRONAFQLRP-FIETRHFLYRGDQERLNPSP 350

287 GALSGRHSPSVGRQHAGCPSTSRPPRPWDTPCPPVYAEKHFYSSGD-KEQLRPSF 345

351 LLSNLQNLTGARLVEIIFLGRPRTSGLCKTHLSRRYQWMPPLFOQLLVNHAQCY 410

346 LLSLRPSLTGARRLVETIFLGRPMWPGTPRLPRLPQRYQWMPPLFLELLGNHAQCPY 405

411 VRLRSHCRFTANQOYTDAL-----NTGPHMLDILLRHSSPPWQV 452

406 GVLLKTCPLRAA---VTPAGVCAREKPGQSVAAPEEDTPRLVLQRLROHSSPWQV 462

453 GFLRACLCVVSASLWTRNRRFPNLIKFTSLGKYKLSLQELMMKMKVEDCHWLRS 512

463 GFVRACLRLVPPGLWGSRRNRRFLNTKFTISLGHAKLSLQELTWKSVRDCAWLR 522

513 SPGRDRVPAEHLREILATELFLWMDTVVOLLRSFVITESTFQKNLFFYKRSVMS 572

523 SPGVGCVPAEHLREILAKFLHMLSVVVELLSFFVYVETITFQKNLFFYKRSVMS 582

573 KLOSIGVROHLRVRRLSQEVRHQDTWMLAMPICRLRFIPKPNGLRPIVNMYSMT 632

583 KLOSIGIQHKLKRVQLRELSEAEVQRREARPAALLSLRFLFKPDGLRPIVNMVYVGA 642

633 RALGRKQAOHQFTORLKTLSMLNYERTKPHLMGSSVLGMDIYRTWEAFVLRVRLDQ 692

643 RTFREKRAERLTSRVKALFSLVNLNERARRPGLLGASVLGLDDIHRWETFLVRVRAQDP 702

QY 693 TPRMTFVADVTGAYDAIPQKLVVEVNVANMIRHSESTYCIROYAVVRDSQGVHKSFR 752  
DB 703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIHK-PQNTYCVRRYAVVQKAHGHVRKAFKS 761  
QY 753 QVTTLSDLQPYMGQFLKHLQSDASALRNSVVIQSIQSMNESSSSLDFDLHFLHRSVVK 812  
DB 762 HVSTLTDLPQYRQFVAHQET--SPLRDAVVIEQSSSLNEASSGLFDVFLRPFMCHEAVR 819  
QY 813 IGRCYTCQCGIPQSSSLSTLCSLCFGDMENKLFQVORDGILLRRFVDDFLVTPHLQ 872  
DB 820 IRKGSYVQCGIPQSSSLSTLCSLCYGDMEKLFAGIRRDGILLRLVDDFLVTPHLTH 879  
QY 873 AKTEFLSTLVHGVPEYGCMLNLOKTVNFPVBPGLGGAAPYQLPAHCLFPWCGLLLDTQT 932  
DB 880 AKTEFLSTLVHGVPEYGCMLNLOKTVNFPVBPGLGGAAPYQLPAHCLFPWCGLLLDTQT 939  
QY 933 LEVFCDSYGAQTSIKTSLTQSVFKAGKTMKLSVLRLKCHGLFLDLQVNSLQVCI 992  
DB 940 LEVQSDYSYARTSIRASLTENRGFKAGNNMRKLFGLRLKCHSLFLDLQVNSLQVCI 999  
QY 993 NYIKFLLQAVRFHACVTLQPDQVVRKNLTFELGIISSOASCCYAILKVNKPGMTLKAS 1052  
DB 1000 NYIKFLLQAVRFHACVTLQPDQVVRKNLTFELGIISSOASCCYAILKVNKPGMTLKAS 1059  
QY 1053 GS---FPPEAAHLCYQAFLLKLAHSHVYKCLLGLPLRTAQKLCRKLPEATMTILKAAA 1109  
DB 1060 GAAGPLPSEAVQWLCHQAFLLKLTETRVYVYVPLLSLRTAQKLCRKLPEATMTILKAAA 1119  
QY 1110 DPALSTDFTILD 1122  
DB 1120 NPALPSDFKTILD 1132

## RESULT 4

US-09-430-323-225

; Sequence 225, Application US/09430323

; Patent No. 6309867

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; Lingner, Joachim

; Nakamura, Toru

; Chapman, Karen B.

; Morin, Gregg B.

; Harley, Calvin

; Andrews, William H.

; TITLE OF INVENTION: No. 6309867el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESSES: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/430,323

; FILING DATE: 29-Oct-1999

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: US 08/724,643



FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-0029300S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
19-430-323-225

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
1 MTRAPCPAVRSLLRSRYREVNPLATFVRRLGPEGRRLVQPGDPKIVRTLVVAQCLVCMHW 60  
1 MPAPRCRAVRSLLRSRYREVNPLATFVRRLGPEGRRLVQPGDPKIVRTLVVAQCLVCMHW 60  
61 GSOPPPADLSFHOVSSKLKELVARVQRLCERNERNVLAFGPELLNEARGGPPMAFTSSVR 120  
61 DARPPPAAPSFQVSCLEKELVARVQRLCERNERNVLAFGPELLNEARGGPPMAFTSSVR 120  
121 SYLPNTVETLRVSGAWMLLSRGVDDLLAVLLAHCALYLLVPPSCAVQVCGSPYQICA 180  
121 SYLPNTVETLRVSGAWMLLSRGVDDLLAVLLAHCALYLLVPPSCAVQVCGSPYQICA 180  
181 TTDWPSVGSASVPTPVGNFTNLRLFLQIKSSRQBEAPKPLALPSRGTNRHLSTLS 240  
181 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPADGARRRGSSASRS 231  
241 VPSAKKARCVPRVREGP-----HRQVLTPSGKSW-VPSAPRSSEVTAEDLSK 292  
232 LPFLPKPRGAPEPRTVVGQGSWAHPGRTGSDRGFCVVSAP-----PABEATSLE 286  
293 GKVSIDLSS-GSVCCCHKFSSTLSLPPQNAPQLRP-PIETRHFLYSRGQDQRLNPSF 350  
287 GALSGRTHSHSPSVGRQHAGPSTSRPPRPMDTPCPVVAETKHFLYSSGD-KEQLRPSF 345  
351 LLSNLQNLGARRLVEILFLGSRPRTSGPLCRTHLSRRVQWMPLEFQQLLVNHAQCY 410  
346 LLSLSPSLTGARRLVETFLGSRPMPGTPRRLPRLPQRYWQWEPLELLGNAQCPY 405  
411 VLLRSHCRPRTANQOVTDAL-----NTSPPHMLDLLLHSSPWQVY 452  
406 GVLLKTHCPILRAA---VTPAAGVCAREKPGQSVAAPEBEDTDFRLVOLLQHSPPWQVY 462  
453 GFRLACLVKVSASLMTGRNRRERFPKLLKFTISLGKYGKLSQLBELMWMKVBDCHMLRS 512  
463 GFVRACLRLLVPPGLMGRNRRERFPKLLKFTISLGKYGKLSQLBELMWMKVBDCHMLRS 522  
513 SPGRKORVPAEHLREILATFLFWLMDTVVQLARSFYITESTFQKNLFFVFKSVWS 572  
523 SPVGVCVPAEHLREILAKFLHMLNMSVVELLRSFFVTTFTFQKNLFFVFKSVWS 582  
573 KLSIGVQRLHRLVRLRELSQBEVRHQDTWLAMPICRLRFIPKPNGLRPIVNMVSGMT 632  
583 KLSIGVQRLHRLVRLRELSQBEVRHQDTWLAMPICRLRFIPKPNGLRPIVNMVSGMT 642  
633 RALGRKQAGHPTQRLKTLFSLMYNERTKHPHLMGSSVLGMNDIYRTWRVFLVRLDQ 692  
643 RTRFRKRAERLTSRKALFVSLNTERARPGLLGASVLGDDIHRAWRTFVLVRAQDP 702  
693 TPRIYFKADVDTGAYDAIPQGLVEVANNIRHSSTYCIROYAVVRDSDQGVKSKFR 752  
703 PPFLYFKVDVDTGAYDTIIPQDLRTVEVIAIIR-PONTYCVRYAVVQKAHGHVKAFS 761

753 QVTTSLDLQPYMGQFLKHLQSDASALRNSVWIEQSIQSSNMVSSSLFDFFLHFLRHSVVK 812  
762 HVSITLTLQPYMRQFVAHLQET--SPLRDAVWIEQSSLINEASSGLFDFVFLRPMCHAVR 819  
813 TGRDYTCQCGIPQSSSLTLLCSLCFGDMENKLFARVQRDGILLRFPVDDFLVTPHLDQ 872  
820 TRGKSYVQCGIPQSSSLTLLCSLCYGDENKLFAGIRRDGILLRFPVDDFLVTPHLDQ 879  
873 AKTFLSTLHVGVPEYGVGMINLOKTVVNFVPEPGTLGGAAPVQLPAHCLFPWCGLLDTOT 932  
880 AKTFLSTLHVGVPEYGVGMINLOKTVVNFVPEDEALGGTAFVQMPAHGLFPWCGLLDTOT 939  
933 LEVFCDSYGYAOTSITKSLTFQSVFKAGTMRNKLJSLVRLKCHGLFLDLQVNSLQTVCI 982  
940 LEVQSDYSYARTSIRASLTFRNGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQTVCT 999  
993 NIYKIFLQAVRFHACVITQLPFDQVRKNTLTFELGIISQASCCYAILKYVNPQMTLKAS 1052  
1000 NIYKILLQAVRFHACVITQLPFDQVRKNTLTFELGIISQASCCYAILKYVNPQMTLKAS 1059  
1053 GS--FPPEAAHMLCYQAFLLKLAHNSVIYKCLIGPLRTAQKLLCRKLPKATMTILKAAA 1109  
1060 GAAGPLPSEAVQMLCHQAFLLKTRHRTVTVYVLLGLSLRTAQKLLCRKLPKATMTILKAAA 1119  
1110 DPALSTDFQTILD 1122  
1120 NFALPSDFKTILD 1132  
RESULT 5  
US-09-128-354-2  
; Sequence 2, Application US/09128354  
; Patent No. 6337200  
; GENERAL INFORMATION:  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants  
; FILE REFERENCE: 015389-003310US  
; CURRENT APPLICATION NUMBER: US/09/128,354  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 08/851,843  
; EARLIER FILING DATE: 1997-05-06  
; EARLIER APPLICATION NUMBER: US 08/854,050  
; EARLIER FILING DATE: 1997-05-09  
; EARLIER APPLICATION NUMBER: US 08/911,312  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: US 08/912,951  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: US 08/915,503  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
; EARLIER FILING DATE: 1997-10-01  
; EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
; EARLIER FILING DATE: 1997-10-01  
; EARLIER APPLICATION NUMBER: US 08/974,549  
; EARLIER FILING DATE: 1997-11-19  
; EARLIER APPLICATION NUMBER: US 08/974,584  
; EARLIER FILING DATE: 1997-11-19  
; EARLIER APPLICATION NUMBER: US 09/052,864  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1132  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-128-354-2  
Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
1 MTRAPCPAVRSLLRSRYREVNPLATFVRRLGPEGRRLVQPGDPKIVRTLVVAQCLVCMHW 60

```

1110 DPALSTDFQTILD 1122
1120 NPALPSDFKTILD 1132

RESULT 6
US-09-675-321-2
; Sequence 2, Application US/09675321
; Patent NO. 6440735
; GENERAL INFORMATION:
; APPLICANT: Gaeta, Federico C.A.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune
; TITLE OF INVENTION: Response to a Telomerase Antigen
; FILE REFERENCE: 015389-003500FC
; CURRENT APPLICATION NUMBER: US/09/675,321
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/112,006
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-675-321-2

Query Match 59.4%; Score 3505; DB 4; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

```

Db 1060 GAAGPLPSEAVQWLCQAEFLKLTTRHYTVYVPLLGSLTAQTQLSRKLPCTTLTALEAAA 1119
Qy 1110 DPALSTDFQTILD 1122
Db 1120 NPALPSDFKTILD 1132

1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGQGRVLRVQGDPAAPRALVAQCLVCPW 60
61 GSQPPADLSFHQVSSKLVARVQCLCERNERNVLAQFELLNEARGPPMAFTSSVR 120
61 DARPPAAPSPFQVSCLELVARVQLCERGAKNVLAQFALLDARGPPPEAFTTSVR 120
121 SYLNTVETLRVSGANWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
121 SYLNTVTDALRGSGAGLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQIGA 180
181 TTDIWPVSASRYRTRPVGRNFTNLRFLQIKSSRQEAQKPLALPSRGTKRHLSTSTS 240
181 ATQARPPPHAS - GPRRLG - - - - - CERANWHSVREAGVPLGLPAPGARRRGGSASRS 231
241 VPSAKKARCYVPRVVEGP - - - - - HRQVLPSPGSKW - VPSPARSPVPTAEKDLSSK 292
232 LPLPRPRRGAAPERTFVQGSWAHPGTRGSDRGFCVSPAR - - - - - PABEATSLE 286
293 GKVDLSLS - GSVCKKHPSSTLSLPPQNAFQLRP - PIETRHFLYSRGQGRINPSF 350
287 GALSCTRSHSPSVGRQHHAGPSTSPRPWDTPCPVYAEVKHFLYSSGD - KEQLRPSF 345
351 LLSNLQNLGARRLVEIFLGSRRPTSGPLCRTHLSRRYQMRPLFOQLLVNHAECY 410
346 LLSLRPSLTGARRLVEIFLGSRRPMPGTPRRLPRLPQRYQMRPLFLELLGNHAQCPY 405
411 VRLRSRCHRTANQQVTDAL - - - - - NTSPPHLMDLLRLHSSPWQYV 452
406 GVLLKTHCPLEAA - - - - - VTPAAGVCAREKPGQSVAAPEEDTDPRLVOLLRQSSPWQYV 462
453 GFLRACLCKVVSASLWGRTHNRERFFKNLKFISLQYKGLSLQELMWKMKVEDCHWLS 512
463 GFVRACLRLVPPGLWGRSHNRERFLNKKFISLQYKGLSLQELTWKMSVRDCAWLAR 522
513 SPGRDVRPAEHLRERILATFLFWMNDTVVQLLRSFFVITESTFQKRLFFYKSVMS 572

```

```

1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGQGRVLRVQGDPAAPRALVAQCLVCPW 60
61 GSQPPADLSFHQVSSKLVARVQCLCERNERNVLAQFELLNEARGPPMAFTSSVR 120
61 DARPPAAPSPFQVSCLELVARVQLCERGAKNVLAQFALLDARGPPPEAFTTSVR 120
121 SYLNTVETLRVSGANWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
121 SYLNTVTDALRGSGAGLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQIGA 180
181 TTDIWPVSASRYRTRPVGRNFTNLRFLQIKSSRQEAQKPLALPSRGTKRHLSTSTS 240
181 ATQARPPPHAS - GPRRLG - - - - - CERANWHSVREAGVPLGLPAPGARRRGGSASRS 231
241 VPSAKKARCYVPRVVEGP - - - - - HRQVLPSPGSKW - VPSPARSPVPTAEKDLSSK 292
232 LPLPRPRRGAAPERTFVQGSWAHPGTRGSDRGFCVSPAR - - - - - PABEATSLE 286
293 GKVDLSLS - GSVCKKHPSSTLSLPPQNAFQLRP - PIETRHFLYSRGQGRINPSF 350
287 GALSCTRSHSPSVGRQHHAGPSTSPRPWDTPCPVYAEVKHFLYSSGD - KEQLRPSF 345
351 LLSNLQNLGARRLVEIFLGSRRPTSGPLCRTHLSRRYQMRPLFOQLLVNHAECY 410
346 LLSLRPSLTGARRLVEIFLGSRRPMPGTPRRLPRLPQRYQMRPLFLELLGNHAQCPY 405
411 VRLRSRCHRTANQQVTDAL - - - - - NTSPPHLMDLLRLHSSPWQYV 452
406 GVLLKTHCPLEAA - - - - - VTPAAGVCAREKPGQSVAAPEEDTDPRLVOLLRQSSPWQYV 462
453 GFLRACLCKVVSASLWGRTHNRERFFKNLKFISLQYKGLSLQELMWKMKVEDCHWLS 512
463 GFVRACLRLVPPGLWGRSHNRERFLNKKFISLQYKGLSLQELTWKMSVRDCAWLAR 522
513 SPGRDVRPAEHLRERILATFLFWMNDTVVQLLRSFFVITESTFQKRLFFYKSVMS 572
523 SPGVGCVPAAEHLRERILATFLFWMNDTVVQLLRSFFVITESTFQKRLFFYKSVMS 582
573 KLSQIGVRQHLERLRELSQEVHRRHOQTWLAMPICRLRPIKPNGLRPIVNMYSMTG 632
583 KLSQIGIRQHLKRVQLRELSAEVROHREAPALLTSRLRPIKPDGLRPIVNMVYVGA 642
633 RALGRKQAOHTQRLKLFMSLNVERTKHPLMGSSVLGNDIYRWRFAVLRVRLAQD 692
643 RTFRREKRAELTSVKALFSLVANYERARRPGLLGASVLGLDDTHRAWRFTVLRVRAQDP 702
693 TPRMTYFVADVTGAYDAIPQGLVEVWAMIRHSESTYICIRQAVVRDQGOVHKSRFR 752
703 PPELYFVKVDVTGAYDTIPODELTEVIAIIK - PONTYCVRYAVVQKAAHGHVRAKFS 761
753 QVTTLSDLQPYMGQFLKHLQSDASALRNSVVIQSTSMNNESSSLPDDFLHFLRHSVK 812
762 HVSTLTDLOPYNRQFVAHQET - - SPLRDVAVIQSSSLNEASSGLFDVFLRMCCHAVR 819
813 IGDRCYTCQGIPOGSSSTLCSLCSFCGDMENKLPFAEVRDGLLRVVDVFLVTHLQD 872
820 IRGKSVQCGGIPQSSSTLCSLCSLCSFCGDMENKLPFAEVRDGLLRVVDVFLVTHLQD 879
873 AKTFTSLTVHGYPCGMINKQTVNPFVEFGTLGGAAQYQLPAHCLFPWCGLLDTQT 932
880 AKTFTRLTVRGVPCVNVLRKTVNPFVEDEALGGTAFVQMPAHLFPWCGLLDTQT 939
933 LEVFCDSYCAOTSITKTSFOSVPKAGTKWENKLLSVLRKCHGLFLDLQVNSLOTVC 992
940 LEVQSDYSSYATSRASITFNRGFAGRNRRKLFVLRKCHSLFLDLQVNSLQVCT 999
993 NIYKIFLQAYRFAHCVQLPDPQVRKXNLTFEFLGIISSQASCCYAILKVNPGMTLKAS 1052
1000 NIYKILLQAYRFAHCVQLPDPHQVQKNNPFFLRVSDTASLQVSLKAKNAGMSLGAK 1059
1053 GS - - - FPPAAHLCVQAFLLKLAHSHVYKCLGLPRTAQKLCRKLPEATMTILKAAA 1109

```

523 SPGVGCVPAABHRLEBILAKFLHMLMSVTVVWELLRSFFVVTSTTFQKRLFFYRKSVNS 582  
573 KLOSIGVROHLERVRRLSQQEEVRRHQDTWLAIPICRLRFIPKPNGLRPIVNMYSYMT 632  
583 KLOSIGVROHLERVRRLSQQEEVRRHQDTWLAIPICRLRFIPKPNGLRPIVNMYSYMT 642  
633 RALGRKQAGHFTORLTKLMSLYBTKPHLMGSSVLGMDIYRTWFAFVLRVRLDQ 692  
643 RTFRREKBAELTSRVKALPSVLYNEARPGSLGASVLGDDIHWAFVLRVRAQDP 702  
693 TPRMYFKADVGTAYDAIPQKIVGVVANIRSESTYICIRQAVVRRDSQQVHKFSRR 752  
703 PPELYFVKVDVTGAYDTIPQDLRTVETIASIK.PQNTYCVRRVAVVQKAAGHVRKAKS 761  
753 QVTWLSLOPMGQFLKHLQSDASALRNSVWIEQISMNNESSLSLDFLHFLRHSVVK 812  
762 HVSTLTDLQPMQFVAHQET--SPURDAVVEQSSLSNEASSGLFVFLRMCCHAVR 819  
813 IGRDCTQCQIGPOGSSLSLTLCSLQCGDMENKLFABVORDGGLLRFVDDFLVTPHLDQ 872  
820 IRGKSVQCCQIGPOGSSLSLTLCSLQCGDMENKLFAGIRRDGGLLRLVDDFLVTPHLDQ 879  
873 AKTELSTLVHGVPEYGCWMLNQLTKTVNFPVEPGTLCGAAPYQLPAHCLFPWCGLLDQT 932  
880 AKTELSTLVHGVPEYGCWMLNQLTKTVNFPVEPGTLCGAAPYQLPAHCLFPWCGLLDQT 939  
933 LEVCDYSGVQTSIKTSLATFQSVKAGTMRNKLJSLVRLKCHGLFLDLQVNSLQTVCI 992  
940 LEVQSDYSSVARTSIRASLTFNFRGKAGRNMRKLFGLVRLKCHGLFLDLQVNSLQTVCI 999  
993 NIYKIFLLQVRFHACVQIFPQDQVRKKNLTFGLIISQASCCYAILKVKVNGPMTLKAS 1052  
1000 NIYKIFLLQVRFHACVQIFPQDQVRKKNLTFGLIISQASCCYAILKVKVNGPMTLKAS 1059  
1053 GS----FPEBAHWLCYQAFLLKLAHNSVTKLLGPIRLTAQKLLCRKLPENWTILKAAA 1109  
1060 GAAGLPSEAVQWLCYQAFLLKLAHNSVTKLLGPIRLTAQKLLCRKLPENWTILKAAA 1119  
1110 DPALSTDFQTLID 1122  
1120 NPALPSDFKTLID 1132

## ULT 7

09-052-919-2

sequence 2, Application US/09052919

atent No. 644650

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Antisense Compositions for Detecting and

TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/052,919

FILING DATE: 31-MAR-1998

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-00360005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-052-919-2

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

Cy 1 MTRAPRCRAVRSLLRSRYREWVPLATFVRIGCGERRLVQCDPKIYRTLVACQLVCMHW 60  
Db 1 MPRAPRCRAVRSLLRSRYREWVPLATFVRIGCGERRLVQCDPKIYRTLVACQLVCMHW 60  
Cy 61 GSOPPPADLSFHCVSSSLKELVARVVRQRCERNERNVLAFFGPELLNEASGPPMAFTSSVR 120  
Db 61 DARPPAAPSFQVSCIKELVARVQRCERAKNVLAFFGPELLDARGGPPPEAFTTSVR 120  
Cy 121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLLAHALCALYLLVPPSCAYQVCGSPLYQICA 180  
Db 121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLLAHALCALYLLVPPSCAYQVCGSPLYQICA 180  
Cy 181 TTDIWFSVSAVSRPTFPVGNFTNLRFQOIKSSSRQEPKPLALPSGTRKHLSTSTS 240  
Db 181 ATQARPPPHAS-GPRRLG-----CERANWHSVRGAVGLPAGARRRGSGSASRS 231  
Cy 241 VPSAKKACVCPVPRVEEGP-----HROVLPTPSKSW-VPSPARSPVETAEKDLSSK 292

```

222 LPLPKRRRGAAPERTPTVQGSNAHPGTRGSPGCVVSPAR-----PABEATSUE 286
293 GKVDLSLS-GSVCKKPKSTSLSPRONAFQARP-FIETRHFLYSRGDOERLNPSF 350
287 GALSOTRHSHPVGRQHAGPSTSRPRPDWTPCPVVAETKHFLYSSGD-KEQLRPSF 345
351 LLSNLOPNLTGARBVLII FLGSRPRTSGPLRTHLSRYWOMRPLFOQLLVNAECQY 410
346 LLSLSRPSLTGARBVLII FLGSRPRTSGPLRTHLSRYWOMRPLFOQLLVNAECQY 405
411 VLLRSHCRFTANQOVTDAI-----NTSPPHLMDLLRLHSSPWQY 452
406 GVLLKTHCPRLAA--VTPAAGVCAREKPGSVAAPBEDTDPRLVQLLRHSSPWQY 462
453 GFLRACLCKVVSASLWTRNRRERFFKNLKFISLGKYGKLSQELMWMKMYEDCHWRS 512
463 GFVRACUERLVPPGLWGRSRNRRERFFKNLKFISLGKHAKLSQELTWKVSVDCAWLR 522
513 SPGKORVPAAEHRLRERILATFLFWMMDTVVQLLRSFFYITESTFQKRLFFYKSVMS 572
523 SPGVCVPAAEHRLRERILAKFLHMLSVVVELLRSFFYITESTFQKRLFFYKSVMS 582
573 KLOSGVQHLERVLRLBELSOEVRHODTWLAMPICRLRFIPKPNGLRPIVNMYSMT 632
583 KLOSGIRQHLKRVQLRELSAEVRCREARPAALLTSRLRFIPKPDGLRPIVNMVYVGA 642
633 RALGRKQAOHFTORLKTFLSMNVYRTKPHLMGSSVLGMNDIYRTWRAFLVRVRLDQ 692
643 RTFREKRAEELTRERKALFVLNFERARPGLLGASVLGLDIDHRAWRTFVLVRQDP 702
693 TPRMYFVKADVTGAYDAIPAOKGLVEVVMNIRHSESTYCIQVAVVRDSQGOVHKSR 752
703 PPELYFVKVDVTGAYDTIPQDLRTVETIASIK-PONTYCVRYAVVQKAAGHVRKAFKS 761
753 QVTTLSLQVPMGOFKHLQSDASALRNSVVBOSISMNESSSLFDFLHFLRSVVK 812
762 HVSTLTDLPQNRQFVAHQET--SPURDAVVLEQSSSLNEASGLFDVFLRFMCHAVR 819
813 LGDRCYTCQGIPOGSSLSLTLCSLCFDMENKLFABEVORDGLLFLVDVDFLLVTPHLDQ 872
820 IRGKSYVQCQGIPOGSSLSLTLCSLCYDMENKLFAGIRRDGLLRLVDDFLVTPHLDH 879
873 AKTFLSTLVHGVPEYGCVMINOKTVNFPVPEGTLGAAAPYQIPAHCLFPWCGLLDDTOT 932
880 AKTFLRTLVRGVPEYGCVMNLRKTVNFPVPEDEALGATFVQMPAHGLFPWCGLLDDTOT 939
933 LEVFCDSYGAQTSIKTSLTPQSVFKAGTKMRNKLKLSVLRKCHGLFLDLQVNSLQTVCI 992
940 LEVQSDYSYARTSIRASLTNRGFKAGNRKLFGLVLRKCHSLFLDLQVNSLQTVCT 999
993 NIYKIFLLOAFRHACVQLPFDORVRKNLTFGLIISQASCCYAILKVNPGWTLKAS 1052
1000 NIYKILLQAVRHACVQLPFPHQWKNPTFFLRVSDTASLCSYLKAKNAGMSLGAK 1059
1053 GS--FPEAAHWCYQAFLLKLAHSHSVYKCLLGPRTAKLLCRKLPEATWTILKAAA 1109
1060 GAAGPLPBEAVQWLCHQAFLLKTHRVTVYVPLLSLRTAQTLQSLRKLPTTILALEAAA 1119
1110 DPALSTDFQTILD 1122
1120 NPALPSDEFTILD 1132

```

SULT 8

-08-912-951-2  
Sequence 2, Application US/08912951  
Patent No. 6475789

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.

```

APPLICANT: Morin, Gregg B.
APPLICANT: Hanley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
TITLE OF INVENTION: THERAPEUTIC METHODS
NUMBER OF SEQUENCES: 335
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,951
FILING DATE: 14-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 435
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00260005
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-912-951-2

```

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

QY 1 MTRAPRCFAVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDPKXYRTLVACQLVCMHW 60
DB 1 MPRAPRCFAVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDPKXYRTLVACQLVCMHW 60
QY 61 GSQPPPADLSFHOVSSIKELVARVQVQLCERNVNLAFGFELLNEARGSGPPMAFTSSVR 120
DB 61 DARPPPAAPSFQVSCLEKELVARVQVQLCERNVNLAFGFELLNEARGSGPPMAFTSSVR 120
QY 121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHLCALYLLVPPSPCAYQVCGSPLYQICA 180
DB 121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHLCALYLLVPPSPCAYQVCGSPLYQICA 180

```

```

181 TTDIWPSVGSAYPTBPVGRNFTNLRFLOQIKSSSRQAPKPLAPSPSGTKRHLSLSTST 240
181 ATQARPPPHAS-GPFRRLG-----CERAWNHSVREAGVPLGLPAPCARRRGCSASRS 231
241 VPSAKKARCVPVPRVEEGP-----HRQVLTPTGKSW-VSPARSPEVPTAEKDLSSK 292
232 LPLPKPRGAAPERTPVQGSWAHPGRTGPDRCVCVSPAR-----PAREATSL 286
293 GKYSIDLSL-GSVCCCHKSSSTLSLPPQNAFQLRP-FIETRHFLYSGRGQGRLNPSF 350
287 GALSGRHSHSPSVGRHAGPSTSPRPWDTPCPVVAETKHFYLSGSD-KEQLRPSF 345
351 LLNLQNLNGARLVEIIFLGRSPTSPLCRTHLSRYWQMPFLQQLLVNHAEOY 410
346 LLSLSPSLTGARLVEIIFLGRSPTSPLCRTHLSRYWQMPFLQQLLVNHAEOY 405
411 VRLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSPWQVY 452
406 GVLLKTECPURAA---VTPAAGVCAREKPGSVAAPEEBDTPRLVQLLRHSGSPWQVY 462
453 GFRLACLVKTVSASLAGTRHNERFPKNLKKETISLGYKGLSLQELVMKVKVEDCHWLS 512
463 GFVRACLRURVPLGMSRHRNRRFLNKKFISLGHAKLSLQELTWKMSVRDCAWLR 522
513 SPGRDVPAAEHLRRLILATFLWLMDTVYVQLLRSFFIITESTFQKNRFFYRKSVM 572
523 SPGVGCVPAEHLRRLILAKFLHLMVSVVVELLSFFYVTTTFQKNRFFYRKSVM 582
573 KLSIGVROHLERVRLAEISQEVRRHQDTWLMPICRLPRTPKNGRLPIVNSYSYGT 632
583 KLSIGVROHLERVRLAEISQEVRRHQDTWLMPICRLPRTPKNGRLPIVNSYSYGT 642
633 RALGRKQAOHQFTORLKTFLSMNLNERTKPHLMGSSVLGMNDIVKTRWAFVLRVRLDQ 692
643 RTFRERKARLTSRVKALFSLVNLNERTKPHLMGSSVLGMNDIVKTRWAFVLRVRLDQ 702
693 TRPMYFVKADTGAYDAIPQKLVUEVVMNRHSSTYCIROYAVVRDSDQGVHKSFR 752
703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIIG-PQNTYCVRRYAVVQKAAHGHVRKAFKS 761
753 QVTLSDLPQMGQFLKHLQSDASALRSVVIQSIQSNSSSLSLDFLHFLHRSVVK 812
762 HVSLIDLPQMGQFLKHLQSDASALRSVVIQSIQSNSSSLSLDFLHFLHRSVVK 819
813 IGRCYTPQCGIPQGSLSLTLCSLCVGMENKLFAGIRRDGLLRVDDFLVTPHLDQ 872
820 IRGKSYVQCGIPQGSLSLTLCSLCVGMENKLFAGIRRDGLLRVDDFLVTPHLDQ 879
873 AKTFLTLVHGVPEYGCMLNLTQVNVFPVPGTGLGAAFPQLPACLPWCGLLDQT 932
880 AKTFLTLVHGVPEYGCMLNLTQVNVFPVPGTGLGAAFPQLPACLPWCGLLDQT 939
933 LEVFCDSYGAQTSIKTSLTFQSVFVAGKATMRNKLKSLVLRKCHGLFLDLQVNSLQTVCI 992
940 LEVFCDSYGAQTSIKTSLTFQSVFVAGKATMRNKLKSLVLRKCHGLFLDLQVNSLQTVCI 999
993 NIYKIFLQYRHFACVQIPLFPQVRKNTLFPGLIISQASCCVAILKXNPGMTLKA 1052
1000 NIYKIFLQYRHFACVQIPLFPQVRKNTLFPGLIISQASCCVAILKXNPGMTLKA 1059
1053 GS---FPPEAAHMLCYOAPLKLAAHSLVYKCLLGLPLRTAQKLLCRKLPEATMTILKAAA 1109
1060 GAAGPLSEAVQMLCHQAFLLKLTTRHVTYVPLLSGLRFTAQQLSRKJPGITLTALAAA 1119
1110 DPALSTDFQITLD 1122
1120 NPALPSDFKITLD 1132

```

ULT 9  
08-974-549A-611  
sequence 611, Application US/08974549A  
atent No. 6166178

```

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974.549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 611:
SEQUENCE CHARACTERISTICS:
LENGTH: 1154 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1154

```

OTHER INFORMATION: /note= "fusion protein composed of hTERT  
OTHER INFORMATION: protein sequence, vector sequences, the  
OTHER INFORMATION: Myc epitope and His6 tag"

08-974-549A-611

1 MTRAPCPAVSLLSRSRYREVMPLATFVRRLGPEGRELVQGPDKIYRILVAOCLVCMFW 60  
1 MPRAPCRVAVSLLSRSRYREVMPLATFVRRLGPEGRELVQGPDKIYRILVAOCLVCMFW 60  
61 GSQPPADLSHVSYSLLKELVAVVQRLCERNERNVLAFLNENARCGPPNAFTSSVR 120  
61 DARPPAPAPSPFRQVSCLELVARVLQRCERGAKNVLAFLNENARCGPPNAFTSSVR 120  
121 SYLPTNVIETLRVSGAMWLLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPTNVTDALRGSGAWGLLRLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASVRFTRPVGRNFTNLRFLQIKSSSROEAPKPLALPSRGTKRHLSTSTS 240  
181 ATQARPPPHAS-GRPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGSGASRS 231  
241 VPSAKKARCPVPRVEEGP-----HRQLPTSPGKSW-VPSPARSPVPTAEKOLSSK 292  
232 LPLPKRGAAPERFVQGSWAHGRGTRGSDGFCVSPAR-----PAEATSLE 286  
293 GKVDLSLS-GSVCKKHPSTSLSPRQNAFQIRP-FIETRHFLYSRGDQGERLNPSF 350  
287 GALSGRHSHPVSGRQHAGPFSRPRPDWTPCPVPVYAEKTHFLYSSGD-KEQLRPSF 345  
351 LLSNLPQNLTGARRLVEIFLGRSPTSGCLTRHLSRRYQWQRPFLQOLLVNAECQY 410  
346 LLSLRPSLTGARRLVEIFLGRSPTSGCLTRHLSRRYQWQRPFLQOLLVNAECQY 405  
411 VLLRSHCRFRFANQCVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
406 GVLLKTHCPRAA---VTPAAGVCAKREKPGQSVAAPEEEDTPRRLVOLLROHSSPWQVY 462  
453 GLRLACLKVVASLWNGTHNRRERFKNKFIKGYGKLSLOELAKMKWCKEDCHWLS 512  
463 GFVACRLRLVPPGWSGRHNRERFKNKFIKGYGKLSLOELAKMKWCKEDCHWLS 522  
513 SPGRKVPAAEHLRERILATFLWIMTYVQLRSFFYITESTFQNRILFFYRKSWS 572  
523 SPGVGCVPAEHLRERILATFLWIMTYVQLRSFFYITESTFQNRILFFYRKSWS 582  
573 KLSIGVROHLERVLRLSLSQEVRRHODTWLAPI CLRPIKPNGLRIVNMSYSGT 632  
583 KLSIGVROHLERVLRLSLSQEVRRHODTWLAPI CLRPIKPNGLRIVNMSYSGT 642  
633 RALGRKQAOHTQRLKTLFSLMNTYRTKHPHMGSSVGLGNDIYRTWRAPVLRVRLDQ 692  
643 RIFPREKRAERLTSRVKALFVLNVERAERFGLGASVGLGDDIHRARVTVLEVRADP 702  
693 TPRIYKADVTGAYDAIPQGLVENVANMHSSESTYCIQYAVVRRDSOGQVHKFR 752  
703 PPLEYKVDVTGAYDTIPQDRLTEVIAIHK-PQNTYCVARYAVQAAHGHVKAFKS 761  
753 QVTTLSDLQPYMGQFKHQDSDASALRNSVVEIOSISMNESSSLPDLFLHFRHSVVK 812  
762 HVSTLTDLPYVWQFVAHQET--SPLRDVAVVIESSSLEASSGLFDVFLRFCHHAVR 819  
813 IGDRCTYCCGIPQSSSLTLLCSLCFGDMENKLPFAVQRDGLLRLFVDDFLVTPHLQ 872  
820 IRGKSVVCCGIPQSSSLTLLCSLCFGDMENKLPFAVQRDGLLRLFVDDFLVTPHL 879  
873 AKTFLSTLVHGVPCVMNLQKTVNPFVEGTGGAPVQLPAHCLFPKCGLLDITQT 932  
880 AKTFLSTLVHGVPCVMNLQKTVNPFVEGTGGAPVQLPAHCLFPKCGLLDITQT 939

QY 933 LEVPCDYSYGAQTSIKTSLTFSQVFKAGTKVKNKLLSVLRKCHGLFELDLQVNSIQTVCI 992  
DB 940 LEVQSDYSYGAQTSIKTSLTFSQVFKAGTKVKNKLLSVLRKCHGLFELDLQVNSIQTVCI 999  
QY 993 NIYKIFELQAYRPHACVIOQLPFDQVRKOLTFEFLGIISQASCCYAILKVKXNPGMTLKAS 1052  
DB 1000 NIYKIFELQAYRPHACVIOQLPFDQVRKOLTFEFLGIISQASCCYAILKVKXNPGMTLKAS 1059  
QY 1053 GS---PPPEAAHWCQAFELKLAHSHVYIKCLGLPLTAQKLLCKLPEATMTTLKAAA 1109  
DB 1060 GAAGPLPSEAVOMLCHQAFELKLTTRHRTVYVFLGSLRTAQOLSKLPFGTTLTALAAA 1119  
QY 1110 DPALSTDFQITILD 1122  
DB 1120 NPALPSDFKITILD 1132

# RESULT 10

US-08-912-951-323  
; Sequence 323, Application US/08912951  
; Patent No. 6475789  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Hartley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
; TITLE OF INVENTION: THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 335  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,951  
; FILING DATE: 14-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002600US  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 323:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

-08-912-951-323

Query Match 59.4%; Score 3505; DB 4; Length 1154;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSILSRVREVPVPLATFVRRIGPSGRRLVQDQPKIVRTLVACQCLVCMHW 60  
1 MTRAPCPAVRSILSRVREVPVPLATFVRRIGPSGRRLVQDQPKIVRTLVACQCLVCMHW 60  
61 GSOPPPADLSFHQVSSLSKELVARVVRQRCERNRNVLAFGPELLNEARGGPPMAFTSSVR 120  
61 DARPPPAAPSFQVCLKELVARVLRQRCERCAKXNVLAFGFALLDGGAGGPPFAFTISVR 120  
121 SYLNPVTIETLVSGAMMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYICA 180  
121 SYLNPVTIETLVSGAMMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYICA 180  
181 TTIDWPSVASYRPTPVGRNFTNLRFQIKSSSQEAPKPLALPSRGTKRHLSTSTS 240  
181 ATQARPAPHAS-GPERRIG-----CERANWHSVREAGVPLGPAPARRRGGASRS 231  
241 VPSAKKACRCYVPRVEEGP-----HRQVLTSPGKSW-VPSPARSPVPTAKDLSSK 232  
232 LPLKPRRGAAPERPPTVPGQSWAHPGRTGSDRGFCVVSAPAR-----PAEAEATSL 286  
293 GKVSDDLSS-GSVCKHKPSSLSLSPQNAFQLRP-FIETRHLYSRGQGOERLNPSP 350  
287 GALSGRHSHPVGRQHAGPPSTSRPRPMDTPCPVPIAETHFLYSSGD-KEQURPSF 345  
351 LLSNLPQNLGARRIVEIFLGSRRPTSGPCRTHLSRRYQWRPLFQQLLVNHAECQY 410  
346 LLSLRPLSLTGARRIVETIFLGSRRPMTGTPRRPLRPQRYQWRPLFELLGNHAQCPY 405  
411 VLLRSHCRFRFANQVTDAL-----NTSPHMLDLRLHSSSPQVY 452  
406 GYLLKTHCPLEAA---VTPAAGVCAREPQGSVAAPDEEDTDRRLVQLLRQHSPPQVY 462  
453 GFLRACLVKWSASLWGTNRHNERFFKNLKFISLGKYKLSLOELMWKMKVEDCHWLS 512  
463 GFVRACLRRLVPPGLWGRHNERPLNTKFIISLGRHAKLSLOELTWKMSVRDCAWLR 522  
513 SPGRORVPAEHRRLERILATFLFWMQTVVQLLRSFFYITESTFQKRLFPYRKSVWS 572  
523 SPGVGCVPAEHRRLERILATFLFWMQTVVQLLRSFFYITESTFQKRLFPYRKSVWS 582  
573 KQSGIGVGRHLRRLRLSQEVEVHQDQWLANPICRLREIPKNGRLRVNMSYSGT 632  
583 KQSGIGVGRHLRRLRLSQEVEVHQDQWLANPICRLREIPKNGRLRVNMSYSGT 642  
633 RALGRKRAQHFTQRLKTLFMSYNTERTKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
643 RTFRREKRAERLTSVKALFSLVINYERARRPGLLGASVLGLDDIHRARWTFVLVRADQP 702  
693 TRMVFYKADVTGADATPOGKLVEVNVANMTHSESYCIROYAVVRDSDGQVHKFSRR 752  
703 PPFLYFVKVDVTGADTTPQDLRLTVIASIIF-PQNTYCYVRYAVVQKAAGHVRKAFKS 761  
753 QVTTLSDLQPMYQGLKLOLSDAALNSVVIQSIQSMNSSLDFFLHFLHRSVVK 812  
762 HVSITLTDLPYMQVQVHAQLOET-SPLDADVIEQSSSLNEASSGLFDVFLRMCWHAVR 819  
813 IGRVCYQCGIPQSSSLSTLCSLCFQGMENKUPAEVQRDGLLRLFRVDDFLVTPHLDQ 872

RESULT 11

US-08-974-549A-613

; Sequence 613, Application US/08974549A

; Patent No. 6166178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin B.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/974,549A

; FILING DATE: 19-NOV-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951



FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 613:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1189  
OTHER INFORMATION: /note= "fusion protein composed of  
OTHER INFORMATION: htrt signal sequence and full length  
OTHER INFORMATION: htrt protein"  
3-08-974-549A-613

Query Match 59.4%; Score 3505; DB 3; Length 1189;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCAVRLSLRLRYREVNPLATVRLGPEGRRLVQDPKLYRVLVAQCLVCMHW 60  
58 MPRAPCRAVRLSLRLRYREVNPLATVRLGPEGRRLVQDPKLYRVLVAQCLVCMHW 117  
61 GSQPPADLSFQVSSKELVARVQVCLERNRNVLAFGLFELLNFAARGGPMATSSVR 120  
118 DARPPAPAPRQVSCVSKELVARVQVCLERNRNVLAFGLFELLNFAARGGPMATSSVR 177  
121 SYLNTVITLRSVCAWMLLSRVGDDLLVYLLAHCALVLLVPPSCAYQVCGSPLYQICA 180  
178 SYLNTVITDALRGSGWGLLLRRVDDVLLHLLARCALFVLVAPSCAYQVCGSPLYQICA 237  
181 TTDVPSVSAVRYRTPVGRNFTNRLFLQIKSSRCQEAQKPLALPSRGTKXHLSTSTS 240  
238 ATQARPPPHAS-GPRRLIG-----CERAWNHSVREAGVPLGLPAPARRRGSGASRS 288  
241 VPSAKKARCYVPRVREBP-----HRQVLPSPGKSM-VPSPARSPVPTAEKDLSSK 292  
289 LPLKGRPRGAAPERTPVQGSWAHPGRTGRPSDRGFCVVSAPAR-----PAEATSLSE 343  
293 GKVDLSLS-GSVCKKHPKPSSTLSLSPRONAQALRP-FIETHFLYSGDQGRINPSF 350  
344 GALSCTHSHPSVGRQHAGPPTSRPPRPMDTCPPEVVAETKHFYSSGD-KEQLRPSF 402  
351 LLSNLQPNLTGARLVELIFLGSRRPTSGPLCRTHRLSRRYWQMRPLFOOLLVNHAEQY 410  
403 LLSLRLPSTGARLVELIFLGSRRPTSGPLCRTHRLSRRYWQMRPLFOOLLVNHAEQY 462  
411 VRLLRSHCRFTANQOYVDAL-----NTSPFHLMLLRHSSPWQY 452  
463 GVLLKTHCPLEAA---VTPAAGVCAKPKQCSVAAPPEEDTDPRLVQLLRHSSPWQY 519  
453 GFLRACLCKVYSASLWGTNRNREFFNKLKFFSLGKYKLSLQELMWKMYVEDCHLRS 512  
520 GFVRACLRLVPPGLWGRSHNRERFLNTRKFFSLGKHAKLSLQELTWKSVRDCAWLR 579  
513 SPGKDRVPAEAHRLRERLATFLFWMMDTVYVQVLLRSFFYITESTTQKNRLLFFYRKSVMS 572

580 SPGVGCVPAEAHRLRERLATFLFWMMDTVYVQVLLRSFFYITESTTQKNRLLFFYRKSVMS 639  
573 KLSQIGVROHLERVELRELSEVVRHQDTWLMAPICRLRFIPKPNGLRPIVNNYSYSGMT 632  
640 KLSQIGIRGHLKRVQLRELSEAEVRQREARPAALLTSRLAFIPKPDGLRPIVNNDYVYVGA 699  
633 RALGRKQAQCHFTORLKTLSMNLNVERTKPHLMGSSVLGNNDIYRTWRAPVLRVLRALDQ 692  
700 RTFRREKRAERLTSRVKALFSLNRYEARREGLLGASVGLGDDIHRARVTFVLRVRAQDP 759  
693 TPRMYFVKADVTGAYDAIPQGLVEVVMIRHSESTYCIQYAVVRRDSQGVYHKSFR 752  
760 PRELYFVKVDTGAYDTIPQDLTEVIAIILK-PQNTYCVRYAVVQCAAHGHVYKAFKS 818  
753 QVTTLSLQPYMGQFLKHLQSDASALRNSVVIQSSISMNESSSLDFLHLFLRHSVVK 812  
819 HVSTLTDLQPYNRQFVAHQET--SPLRDVAVIEQSSSLNEASSGLFDVFLRFMCHHAVR 876  
813 IGRCYTQCOGIPQGSISLTLCSLCFGDMENKLFAYQORDGLLLRFVDDFLVTPHLDQ 872  
877 IRGKSVQCOGIPQGSISLTLCSLCYGMENKLFAGIRRDGLLLRLVDDFLVTPHLDQ 936  
873 AKTFLSTLVHGVPEYGCMLNOKTVVNFPEGLTGAAPYQLPAHCLFPWCGLLLDTOT 932  
937 AKTFLSTLVHGVPEYGCMLNOKTVVNFPEGLTGAAPYQLPAHCLFPWCGLLLDTOT 996  
933 LEVFCDSYGAOTSITKSLTQSFVFKAGTWRNKLTVLRKCHGLFLDLQVNSLQTVCI 992  
997 LEVQSDYSYARTSIRASLTFRGFKAGNRWRKLVGLVRLKCHSLFLDLQVNSLQTVCT 1056  
993 NIYKIFLQAYRFHACVLIQLPDQVRNKLTFGLITSSQASCCYAILKVKXPGMTLKAS 1052  
1057 NIYKILLQAYRFHACVLIQLPDQVRNKLTFGLITSSQASCCYAILKVKXPGMTLKAS 1116  
1053 GS---FPPAAHWLCVQAFLLKLAHSLVYKCLIGPLRTACKLCKRLPEATMTILKAAA 1109  
1117 GAAGPLPSEAVOMLCHQAFELKLTNRVTVYVPLGSLRTAQTQLSRKLPGTILTALEAAA 1176  
1110 DPALSTDFQITILD 1122  
1177 NPALPSDFKTLID 1189

RESULT 12  
US-08-912-951-325  
Sequence 325, Application US/08912951  
Patent No. 6475789  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Liringer, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Hartley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
NUMBER OF SEQUENCES: 335  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951

FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 325:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
08-912-951-325

Query Match 59.4%; Score 3505; DB 4; Length 1189;  
est Local Similarity 62.4%; Pred.No. 0;  
atches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLLRSRYREYVWPLATVRLGPEGRLVQPGPKIYRILVAQCLVCHW 60  
58 MTRAPCPRAVSLRSRYREYVWPLATVRLGPGWELVQVGPAAFRALVAQCLVCEW 117  
61 GSQPPADLSFHOVSSLLKELVARVVOQLCERNERNVLAQFELLNEARGPPMAFTSSVR 120  
118 DARPPPAAPSPFQVSCUKELVARVVOQLCERNERNVLAQFELLNEARGPPMAFTSSVR 177  
121 SYLNTVETILRVSGAMWLLLSRVGDDLLVLLAHCAALYLLVPPSCAYQVCGSPYQICA 180  
178 SYLNTVETILRVSGAMWLLLSRVGDDLLVLLAHCAALYLLVPPSCAYQVCGSPYQICA 237  
181 TTDIWPVSASVYRTPVGRNFTNLRFLOQIKSSRQEAQKPLALPSRGTKRHLSTSTS 240  
238 ATQARPPPHAS-GRPRUG-----CERAWNSVREAGVPLGLFAPGARRGGASAS 288  
241 VPSAKKACVYVPRVEEGP-----HRQVLPFGSKW-VPSPARSPVPTAEKDLSSK 292  
289 LPLPKRPRGAAPPEPTEPVQGGSWAHPGRTGFSDRGFCVSPAR-----PAEAETSLE 343  
293 GKVSIDLIS-CSVCKKHPSTLSLSPRQNAQQLRP-FLETBFLYSRGGDGOERLNPSF 350  
344 GALSCTRHSFVSQRHAGPPSTSRPRPMDTFCPEVYAEKTHFLYSSGD-KEQLRPSF 402  
351 LLSNLQPNLTGARRLVEIFLFGSPRTSGPLCRTHLSRRYQWVRPLFQQLLVNHAECY 410  
403 LLSLRPSLTGARRLVEIFLFGSPRTSGPLCRTHLSRRYQWVRPLFQQLLVNHAECY 462  
411 VLLRSHCRFRFANQVTDAL-----NTSPHMLDILLRHSSPWQY 452  
463 GVLLKTHCFPLRAA---VTPAAGVCAKPGQSVAAPEEDTDPRLVQLLRQHSPPWQY 519

QY 453 GFLRACLCKVASLWGTNRHNERFFKNLKKFISLGGYKGLSLOELAWMKVEDCHWLS 512  
DB 520 GFVRACLRRLVPPGLWGRHNERRLNRTKKFISLGGHAKUSLOELTWKSVRDCALRR 579  
QY 513 SPQKORVPAAEHRLRERILATFLWIMDTYVQVLLRSFFYITESTFQKRLFFYKRSVWS 572  
DB 580 SPQVGVCPAAEHLRERILAKFLHLMVSVVVELLRSPFYVTTTFQKRLFFYKRSVWS 639  
QY 573 KLOSIGVROHLERVLRELQBEVRRHODTWLAPICELRIKPKNGRLRPVNMYSWGT 632  
DB 640 KLOSIGIRQLHRLVQLRELSEABVQREARPAALLTSRLRFIPKPDGLRPVNMVYVGA 699  
QY 633 RALGRKQOACHFTQRLKTLFSLMYNERTKHPHMGSSVLGMNDIYRTWRAFLVLRALDQ 692  
DB 700 RTRERREKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRAWRTFLVLRADQ 759  
QY 693 TPRMYFVKADVTGAYDAIPOGKLVEVVMNMRHSESTYCIROYAVVRADSDGOVHKSPFR 752  
DB 760 PPELYFVKVDVTGAYDTIPQDLTEVIASIIK-PQNTYCVRRYAVVQKAAHGHVYKAPKS 818  
QY 753 QVTTLSLOPYMGQFLKHLQDSASALRNSVVIEOSISMNNESSSLDFDFLHFLRHSVVK 812  
DB 819 HVSTLTLQPYRQFVAHLQET--SPLRDVAVVIEOSSSLNEASSGLDFVFLRFCHAVR 876  
QY 813 IGRCTYTCQCGIPQSSSLSTLCSLCFGDMENKLPFAVQORDGLLLRFVDDFLVTPHLDQ 872  
DB 877 IRCKSVYQCGIPQSSSLSTLCSLCYGMENKLPFAGIRRDGLLLRLVDDFLVTPHLDQ 936  
QY 873 AKTFLSLVHGVPEYCGMINLOKTVNRPVPEPGLTGGAPYQLPAHCLFPMWGLLDTQT 932  
DB 937 AKTFLSLVHGVPEYCGMINLOKTVNRPVPEPGLTGGAPYQLPAHCLFPMWGLLDTQT 996  
QY 933 LEVFCYSGVAOTS-KTSLTFQSVFKAGTKMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
DB 997 LEVQSDYSSVARTSIRASLTFRNGFKAGNWERKLFVLRKCHSLFLDLQVNSLQTVCI 1056  
QY 993 NIYKIFLQAYRPHACVIOQLPDQVRKNLIFLGLIISOQASCCYAILKVKNGMTLXAS 1052  
DB 1057 NIYKIFLQAYRPHACVIOQLPDQVRKNLIFLGLIISOQASCCYAILKVKNGMTLXAS 1116  
QY 1053 GS---PPPEAAHWLCYQAFLLKLAHSAVYKCLLQPLRTAQKLLCRKLPPEATMTILKAAA 1109  
DB 1117 GAAGPLPSEAVQWLCQAFLLKTHRTVTVYVPLLSLRTAQKLLCRKLPPEATMTILKAAA 1176  
QY 1110 DPALSTDFOTILD 1122  
DB 1177 NPALPSDFKTILD 1189

## RESULT 13

US-08-974-549A-612  
Sequence 612, Application US/08974549A  
Patent No. 6166178

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Hatley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 612:

SEQUENCE CHARACTERISTICS:

LENGTH: 1200 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..1200

OTHER INFORMATION: /note= "fusion protein composed of His6

OTHER INFORMATION: and Anti-Xpress tags, enterokinase

OTHER INFORMATION: cleavage site and full length hprt

OTHER INFORMATION: protein"

i-08-974-549A-612

Query Match 59.4%; Score 3505; DB 3; Length 1200;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVRSLLRSYREVWPLATFVRLGEGRLVQDPDKIYRTLVAAQCLVCMHW 60

69 MPRAPCRAVRSLLRSYREVWPLATFVRLGEGRLVQDPDKIYRTLVAAQCLVCMHW 128

61 GSOPPPADLSFHQVSSLKELVARVQRLCERNERNVLAFCFELLNBARCGPPMAFTSSVR 120

129 DARPPPAPEFRQVSCLELVARVLQELCERGAKNVLAFCFALLDGCARGPPPEATTSSR 188

121 SYLNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180

189 SYLNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 248  
181 TTDIWPVSASVYPTREVGRNTNRLFLQOIKSSSSROEAPKPLALPSRTGKXHLSTSTS 240  
249 ATQARPPPHAS-GPRRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGSASRS 299  
241 VPSAKKARCYPVPRVEEGP-----HRQVLTPSGKSW-VSPASPSPEVPTAEKDLSSK 292  
300 LPLPKRRRGAAPPEPPTVQGSNAHPGRTGSPDRGFCVVSFAR-----PABEATSL 354  
293 KGVSDLSLS-GSVCKHKPSSTSLSPRONAPQLRP-FIETRHFLYSRGDQGERLNPSF 350  
355 GALSOTRHSHPVGRQHAGPPSTSRPPREWDTPCPVVAETKHFLYSSGD-KEQLRPSF 413  
351 LLSNLQPNLIGARBLVEIILFGSRPRTSGPLCTHLSRRYQWMLFOQLLVNHAECY 410  
414 LLSLRPSLTGARELVETIFLGRPMWPGTPRLPRLPQRYQWMLFLELGNHAQCXY 473  
411 VLLRSHCRPTANQQVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
474 GVLLKTHCPLEAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRLLVQLLRQHSSPWQVY 530  
453 GFLRACLCVVSASLNGTRNERRFFXNKKFISLGYKLSLOELMWKMKVEDDCHWLAS 512  
531 GFVRACLRRLVPPGLNGSRNERRFLENTKFKISLGHAKLSLOELTWKSVSRDCAWLR 590  
513 SPGKDRVPAABHRLRERILATFLFWMMDTVVOLLRSFFYITESTFQKNRLFYKRSVMS 572  
591 SPGVGCVPAABHRLRERILAKFLHLSMVVVELLSRFFVYVTTTQKNRLFYKRSVMS 650  
573 KLSIGVFOHLRLVRLRELSQBEVRRHODTWLAMPICRIARFIPKPNGLRPIVNMVSMGT 632  
651 KLSIGIRQHLKRLVRLRELSQBEVRRHODTWLAMPICRIARFIPKPNGLRPIVNMVSMGT 710  
633 RALGRKQAOHFTORLKTLSMLANVETKHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
711 RTFREKRAERLTSRVKALPSVLNTERARPGLLGASVLGLDDIHRANRTFVLVRAQDP 770  
693 TPRMYFVKADVTGAYDAIPQGLVEVVMNIRHSESTYCIQYAVVVRDQSGQVHKSFR 752  
771 PPELYFVKVDTGAYDTIPQRLTEVIASIIK-PQNTYCVRYAVVQKAAHGHVRKAFS 829  
753 QVTTLSDLQPYMGFLKHLSDSASALRNSVVIQSIQSMNESSSLSLDFDFLHFLRHSVK 812  
830 HVSITLTDLPYMRQFVAHLQET--SPLRDVAVIEQSSSLEASGLFDVFLRFCHAVR 887  
813 IGDRCYTQCGIPQGSLSLTLCSLCFQDMENKLFQVORDGLLLRVDDFLVTPHLQ 872  
888 IRGKSYVQCGIPQGSLSLTLCSLCYQDMENKLFAGIRRDGLLLRVDDFLVTPHLTH 947  
873 AKTFLSTLVHGVPEYGCMINLQKTVNPFPEPTLGGAPYQLPAHCLFPWCGLLDTOT 932  
948 AKTFLRTLVRGVPEYGCVINLRKTVNPFPEDEALGTAFAVQMPAHGLFPWCGLLDT 1007  
933 LEVFCDSYGAQTSIKTSITQSVFKAGKMTWENKLLSVLRKLGHLFLLQVNSLQTVCI 992  
1008 LEVQSDYSYARTSIRASITNRCFKAGNRRKLFGLVLRKLGHLFLLQVNSLQTVCT 1067  
993 NIYKIFLQAYRFHACVQLQPDQVRKXNLTFFLGITISSQASCCYAILKYNPQMTLKAS 1052  
1068 NIYKILLQAYRFHACVQLQPFHQQWKNPTFFLRVSDTASLCYSILKAKNAGMSIGAK 1127  
1053 GS---PPPEAAHLCVQAFLLKLAHNSVIYKCLLGLRTAQKLCRKLPRATMTILKAAA 1109  
1128 GAAGPLPSEAVNWLCHQAFLLKLTREHVTYVPLIGSLFETQQLSRKLPCTTLTALEAAA 1187  
1110 DPALSTDFQITLD 1122  
1188 NPALPSDFXTILD 1200

RESULT 14

61	GSOPPNADLSFHQVSSIKELVARVQORLCERNERNVLA7GPELLN2ARGGPPNAPFTSSVR	120
20	DAIP6PAAPAFROVSCLELVARVQORLCERGAKNVLA7GALLDUGARGPPNAPFTSSVR	138
21	SYLPNTVIETLRVSGAWMLLSRVGDDLVYLLAHACALYLLVPPSCAYQVCGSPLQICA	180
89	SYLPNTVTDALRGSGAWGILLRRVGDVVLVHLLARCALFVLVAPSCAYQVCGPPLVOLGA	248
81	TTDIPSVSASVPTPEVGENFTNLRFLOQIKSSRSQEA6PKPIALPSRGTKHLSLTSTS	240
449	ATQARPPHAS-GPRRLG-----CERAMNHSVREACVPLCLFAPGARRRGGSASRS	299
441	VPSAKKARCAYVPRVBEGP-----HRQVLP7PSQKSW-VPSPARSPVPTAEKDLUSK	292
400	LPLPKPRRGAAPERTPVQGSWAHPGRTGRPSDRGFCVSPAR-----PABEATSLE	354
493	GKVSIDLSS-GSVCCKHKSSTSLAPPRONAFQLRP-FLET2HFVLSRGDGERLNPSF	350
555	GALSCTRHSHPSVGRQHAGPPTSPRPWDTPCPVVAETHKFLYSSGD-KEQLPSPF	413
51	LLSNLOPNLTGARRLVEIIFLGSRPRTSGPLCTHLSRRYQWGRPLFOOLLVNHAEQY	410
114	LLSSLSPSLTGARRLVETIFLGSRPMPGTPRPLPRLPQRYQWGRPLFLELLGNHACPY	473
111	VRLRSHCRERTANQVTDAL-----NTSPHMLDLRLHSSPMQVY	452
74	GVLKTHCPURAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPMQVY	530
53	GFIRACLCKVVSASLMGTRHNERFFKNLKKF7SLQYKGKLSLOELMKWKV3DDCHWLSR	512
331	GFVRACLRRLVPPCLMGSRHNERFRLNTKKF7SLQGHAKLSLOELTWKMSVRDCAWLRR	590
113	SPKORVPAAEHLREILATLFLWMDTVVOLLSSFFVITESTQKNRLFYRKSVWS	572
91	SPGVGCVPAEHLREILAKFLHNLMSVYVVELLSSFFVITET7QKNRLFYRKSVWS	650
73	KLOSIGVROHLRVLRELSQEEVRHHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMGT	632
51	KLOSIGIRQHLKTVQLELSEAEVROHREARPA7LTSRLRFIPKPDGLRPIVNDYVYVGA	710
333	RALGRKQAOHF7ORLKTLPFMLNYERTKPHLMGSSVLGMNDIYRTWRAFVLVRALDQ	692
111	RTFRKRAEELTSRVKALFVSNLYEARPGLLGASV7GLDDIHAWR7FVLEVRAQDP	770
493	TPRYFVKADVTGAYDAIPGKLVVVMNMRHSESTYCI7QYAVVRDRSQGVHKSFRR	752
71	PPELYFVKVDVTGAYDTIPQDRUTEVIA7SIK-PONTYCVRRYAVVQKAHGHVRKAFKS	829
53	QVTTLSLQPYMGQFLKHLQSDASALRNSV7IEQSI3MNESSS7LFDPLFLRLHSVVK	812
30	HVSTLTDLPYMQFQVANHQET--SPLRDAV7IEQSSSLNEASSG7FDVFLRPMCHHAVR	887
113	IGDRCYTCQCGIPQGSLSL7LLCSLFCGDMENKLP7AEVQRDGLLLRVPD7FLV7LTHLDQ	872
88	IRGKSVYQCGIPQGS7IL7LLCSL7CGVDMENKLP7AGIR7DGLLLR7LVD7FLV7LTH	947
73	AKTFLSTLVHGVEPYGCMINLQKTVN7FVPEG7LGGAA7VQLPAHCL7PWCGLLLDTQ	932
448	AKTFLRTLVRGVEPYGCVNLRKTVN7FV7EDSALG7G7AFVQ7PAHGL7PWCGLLLDTR	1007
333	LEVPCDYSYGAQTSIK7SL7P7SQSVFKAGK7WRNKL7SLV7LCK7HGL7PLD7QVNSLQTVCI	992
008	LEVQSDYSYARTS7IASL7P7FRNGFKAGRNNR7KL7FQV7L7K7CHSL7FLD7QVNSLQTVCT	1067
493	NIYKIFLLOQVR7HACV7QLP7EORV7KNLT7F7LGI7SSQA7CCY7AIL7KVN7PGMTLKAS	1052
468	NIYKILLQV7R7HACV7QLP7HQQW7KNT7F7L7V7SD7AS7CY7SIL7KAK7NAG7SL7GAK	1127
553	GS---FPPEAAHMLCVQAF7LLK7LA7HSV7IYK7LLG7PL7R7AQ7L7CK7PL7EAT7WTL7KAAA	1109
128	GAAGP7PSEAVQML7CHQAF7KLT7HR7V7V7PL7LS7R7AQ7L7SK7L7P7TTL7ALE7AAA	1187

1110 DPALSTDFQILD 1122  
: : : : :  
1188 NPALPSDFKILD 1200

## RESULT 15

08-974-549A-600  
Sequence 600, Application US/08974549A  
Patent No. 6166178

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 600:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1285 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1285  
OTHER INFORMATION: /note= "fusion protein composed of  
OTHER INFORMATION: enterokinase cleavable, His tagged  
OTHER INFORMATION: thioredoxin moiety and full length hprt"

US-08-974-549A-600

Query Match 59.4%; Score 3505; DB 3; Length 1285;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY	1	MTRAPRCPAVRSLRSRYEVNPLATFVRRLLGPEGRRLVQDPFKIYRTLVLAQCLVCMHW	60
DB	154	MPAPRCRAVRSLRSRYEVNPLATFVRRLLGPEGRRLVQDPFKIYRTLVLAQCLVCMHW	213
QY	61	GSOPPPADLSFHOVSLKELVARVQRLCERNERNVLAFFELNEARGGPPMAFTSSVR	120
DB	214	DARPPPAAPSFQVSCLELVARVQRLCERNERNVLAFFELNEARGGPPMAFTSSVR	273
QY	121	SYLPNTVITLTVSGAMLLSRVGGDLLVLLAHALCALYLLVPPSCAYQVCGSPLYQICA	180
DB	274	SYLPNTVITLTVSGAMLLSRVGGDLLVLLAHALCALYLLVPPSCAYQVCGSPLYQICA	333
QY	181	TTDIWPSVASRYRTPRGVGRNFTNLRFCQIKSSRQBPAPLALPSRGTGRHLSLSTS	240
DB	334	ATQARPPPHAS-GPARRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS	384
QY	241	VPSAKZARCYVPVVEEGP-----HRQVLTPTGKSW-VPSPARSPVETAEDKLSKK	292
DB	385	LPLPKPRRGAAPERTPVGGGSAHPTGRTGPDGFCVVSAR-----PAEATSL	439
QY	293	GKVSDELSS-GSVCKKHPSSTLSPPRONAFQRP-FIETRHLYSRGQGRRLNPSF	350
DB	440	GALSGTRHSHPSVGRQHAGPPSTSRPPRMDTPCPVVAETKHPFLYSSGD-KEQLRPSF	498
QY	351	LLSNLQPNLTGARRIVEIIFLGRSPTSGPLCRTHLSRYWQMRPLFQQLLVNHAQCY	410
DB	499	LLSSLRPSLTGARRIVETIFLGRSPTSGPLCRTHLSRYWQMRPLFQQLLVNHAQCY	558
QY	411	VRLLRSHCRFTANQOVIDAL-----NTSPHMLDLRLHSHSPQVY	452
DB	559	GVLLKTHCPILRAA---VTPAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPQVY	615
QY	453	GFLRACLCKVVSASLWGTNRNERRFKLKKFISLGKYKLSLQELMKMKVDECHLRS	512
DB	616	GFVRACLRLVPELWGSRRNRRFLRNTKFFISLGKAKLSLQELMKMKVDECHLRS	675
QY	513	SPGKDRVPAAEHRLRERILATFLFWMIDTVVQLRSFFYITESTFQKNLFFYRKSWS	572
DB	676	SPGVCVPAAEHRLRERILATFLFWMIDTVVQLRSFFYITESTFQKNLFFYRKSWS	735
QY	573	KLOSIGVROHLEVRRLRELSEVREHHDQTLAMPICRIRIPKPNGLRPIVNMYSYMT	632
DB	736	KLOSIGVROHLEVRRLRELSEVREHHDQTLAMPICRIRIPKPNGLRPIVNMYSYMT	795
QY	633	RALGRBKQAHFTQRLKTLPSMLNVERTKHPHMGSSVLGMDIVRTWRAFLVRALDQ	692
DB	796	RTPRRKRERLTSRKALPSVNLVERARRPGLLAGSLGLDDIHEANRTFVLVRAQDP	855
QY	693	TPRMFYKADVTGAYDAIPQGLVEVANNMHSSTYICRQYAVVRDSCQVHKSPFR	752
DB	856	PELYFKVDVTGAYDTIPQDRLTEVIASIIK-PQNTYCVRYAVVQKAAHGHVRKAFKS	914
QY	753	QVTTLSLQPYMCOFLKHLQDSNASLRNSVIEGSIWNTESSSLPDLFELHLSHVVX	812
DB	915	HVSTLTDLQPYMCOFLKHLQDSNASLRNSVIEGSIWNTESSSLPDLFELHLSHVVX	972



GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.  
 protein - protein search, using sw model  
 on: February 4, 2004, 14:11:55 ; Search time 46 Seconds  
 (without alignments)  
 5107.107 Million cell updates/sec

tle: US-09-042-460-2  
 rfect score: 5901  
 quence: 1 MTRAPRCRAVRLSRVRE.....TILKAADPALSTDFQTILD 1122

ring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

arched: 801455 seqs, 209382283 residues  
 tal number of hits satisfying chosen parameters: 801455

nimum DB seq length: 0  
 ximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
 Listing first 45 summaries

atabase : Published Applications AA:\*  
 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.\*  
 2: /cgn2\_6/ptodata/1/pubpaa/PTC\_NEW\_PUB.pdb.\*  
 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.\*  
 4: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*  
 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.\*  
 6: /cgn2\_6/ptodata/1/pubpaa/PTCUS\_PUBCOMB.pdb.\*  
 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.\*  
 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*  
 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pdb.\*  
 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pdb.\*  
 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pdb.\*  
 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.\*  
 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pdb.\*  
 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pdb.\*  
 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pdb.\*  
 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pdb.\*  
 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.\*  
 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	4751	80.5	1128	12	US-10-295-681-53
2	3505	59.4	1132	10	US-09-990-080-2
3	3505	59.4	1132	10	US-09-749-7288-31
4	3505	59.4	1132	10	US-09-843-676-225
5	3505	59.4	1132	10	US-09-953-052-2
6	3505	59.4	1132	12	US-10-295-681-57
7	3505	59.4	1132	15	US-10-053-758-225
8	3505	59.4	1132	15	US-10-208-243-2
9	3505	59.4	1132	15	US-10-054-295-225
10	3505	59.4	1132	15	US-10-054-611-225
11	3505	59.4	1132	15	US-10-105-363-2
12	3505	59.4	1132	15	US-10-044-692-2
13	3505	59.4	1132	15	US-10-044-539-2
14	3505	59.4	1154	15	US-10-044-692-323
15	3505	59.4	1154	15	US-10-044-539-323

16	3505	59.4	1189	15	US-10-044-692-325	Sequence 325, App
17	3505	59.4	1189	15	US-10-044-539-325	Sequence 325, App
18	3505	59.4	1200	15	US-10-044-692-324	Sequence 324, App
19	3505	59.4	1200	15	US-10-044-539-324	Sequence 324, App
20	3505	59.4	1285	15	US-10-044-692-314	Sequence 314, App
21	3505	59.4	1285	15	US-10-044-539-314	Sequence 314, App
22	3498	59.3	1132	12	US-10-385-882-2	Sequence 2, Appli
23	3496	59.2	1407	15	US-10-044-692-334	Sequence 334, App
24	3496	59.2	1407	15	US-10-044-539-334	Sequence 334, App
25	2492.5	42.2	1003	10	US-09-843-676-217	Sequence 217, App
26	2492.5	42.2	1003	11	US-09-438-486-217	Sequence 217, App
27	2492.5	42.2	1003	15	US-10-053-758-217	Sequence 217, App
28	2492.5	42.2	1003	15	US-10-054-295-217	Sequence 217, App
29	2492.5	42.2	1003	15	US-10-054-611-217	Sequence 217, App
30	2355	39.9	1131	12	US-10-295-681-52	Sequence 52, Appli
31	2195	37.2	807	15	US-10-044-692-5	Sequence 5, Appli
32	2195	37.2	807	15	US-10-044-539-5	Sequence 5, Appli
33	2132	36.1	622	15	US-10-294-778-12	Sequence 12, Appli
34	1825.5	30.9	500	12	US-10-282-960-81	Sequence 81, Appli
35	1686.5	28.6	364	12	US-10-304-095-41	Sequence 41, Appli
36	1511.5	25.6	438	15	US-10-294-778-10	Sequence 10, Appli
37	1506.5	25.5	564	10	US-09-843-676-101	Sequence 101, App
38	1506.5	25.5	564	10	US-09-766-253-101	Sequence 101, App
39	1506.5	25.5	564	11	US-09-438-486-101	Sequence 101, App
40	1506.5	25.5	564	15	US-10-053-758-101	Sequence 101, App
41	1506.5	25.5	564	15	US-10-054-295-101	Sequence 101, App
42	1506.5	25.5	564	15	US-10-054-611-101	Sequence 101, App
43	1497	25.4	437	15	US-10-294-778-2	Sequence 2, Appli
44	1249.5	21.2	364	12	US-10-304-095-40	Sequence 40, Appli
45	1105.5	18.7	538	15	US-10-044-692-316	Sequence 316, App

ALIGNMENTS

RESULT 1  
 US-10-295-681-53  
 ; Sequence 53, Application US/10295681  
 ; Publication No. US2003016270A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: E. Premkumar Reddy  
 ; APPLICANT: Sushil G. Rane  
 ; APPLICANT: Richard V. Mettous  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REVERSIBLY  
 ; INDUCING CONTINUAL GROWTH IN NORMAL CELLS  
 ; FILE REFERENCE: 6056-307  
 ; CURRENT APPLICATION NUMBER: US/10/295,681  
 ; CURRENT FILING DATE: 2002-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/334,760  
 ; PRIOR FILING DATE: 2001-11-15  
 ; NUMBER OF SEQ ID NOS: 69  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 53  
 ; LENGTH: 1128  
 ; TYPE: PRT  
 ; ORGANISM: Mesocricetus auratus  
 US-10-295-681-53

Query Match	80.5%	Score 4751;	DB 12;	Length 1128;
Best Local Similarity	80.1%	Pred. No. 0;		
Matches 304;	Conservative	90;	Mismatches 127;	Indels 8;
Gaps				4;
Qy	1	MTRAPRCRAVRLSRVRE	PLATFVRRLGSEGRRLVQPGDPKIVRLVAOCLVCMFW	60
Db	1	MTRAPRCRAVRLSRVRE	PLATFVRRLGSEGRRLVQPGDPKIVRLVAOCLVCMFW	60
Qy	61	GSQPPADLSPHVSSLSKELVARVQRLCERNRNVLAFQFELLNEARGGPPMFAFTSSVR	120	
Db	61	DSQPPADLSPHVSSLSKELVARVQRLCERNRNVLAFQFELLNEARGGPPMFAFTSSVR	120	
Qy	121	SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLALHLCALYLLVPPSCAYQVCGSPLYQICA	180	
Db	121	SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLALHLCALYLLVPPSCAYQVCGSPLYQICA	180	



181 TTDIWPVSASYPRTVRVGNFNNLPLQOIKSSSQEAPKPLALPSRGTCKHLSLTS 240  
181 TAETWPSVSIYRPTREVGNGFTHLGSSTHRSNQHAEWKPPPLPSREAKSLSTVRS 240  
241 VPSAKKARCVFVREBEGPRQVLPTPSGKSWPSPARSPEVP---TAEDKLSGKVS 297  
241 VPPSKKARCDLAPLEKGPYQAVPTPSDKTWPNPAKSHAVPISTTKEDLSSGVKAP 300  
298 LSLSGSCCKHKSSTLSLSPQONAPQLPFTIETHLYSRGDOGERNLPSFLSNLQ 357  
301 LSRGSCVYKHKPSSTLSQSPQLQONAPQLPFTIETHLYSRGGRERLNPFLNQLQ 360  
358 NLTCARLRIETIFLGSRPRTSGPLCRTHLSRRYQWRPLFOQLLVNHAECYVLLRSH 417  
361 SLTGARLVELFLQMPRTSGPLCGRRRLSKYQWRPLFOQLLVNHAECYVLLRSH 420  
418 CRFTANQOQTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCKVKVSASLWGTNR 476  
421 CRFTAAHQVAGALNTTSPQRLMNLRLHSSPMQVYGFQLQACVGLVPPLWGSRNQR 480  
477 FPKNLKXIFSLGKYSLOELMKMKVEDCHLRSRSPGKORVPAEAHRLRILATLF 536  
481 FFKVXKPFISLGYDKLSLQELTWKVKQDCRWLRSSPGNCPVAAEHRTREILAVLF 540  
537 WLMQTYVQLLSRSPFYITESTFQKRLFFYRKSVWSKLSIGVQCHLRLRLSSEV 596  
541 WLMQTYVQLLSRSPFYITESTFQKRLFFYRKSVWSRQLSIGVQCHLRLRLSSEV 600  
597 RHQDWTLMPICLRLIPKPNGLRPIVNNYSNGTRALGRKQAOHFTQRLTFSMLN 656  
601 RORQEAWFAMPICRLRIPKPSGLRPIVNNYS-NGTRAFDKGQAOHFTQRLTFSMLN 659  
657 YERTKHPHMGSSVGLNDIYRTWRAVLAVALDQTPRMVFKADVTGAYDAI PQGLV 716  
660 YELTKHNLGASVGLNDIYRTWRAVLAVALDQTPRMVFKADVTGAYDAI PQGLV 719  
717 EVANMIRHSESTYCIQYAVVRDSQGVHKSFRQVTTLSLQPYMGQFLKHLQSDA 776  
720 EVANMIRHSESTYCIQYAVVRDSQGVHKSFRQVTTLSLQPYMGQFLKHLQSDA 779  
777 SALRNSVVIQSIIMNBSLSDFLFLHLSVVKIGRKYTCOCGIPQSGSLSLTLC 836  
780 SALRNSVVIQSIIMNBSLSDFLFLHLSVVKIGRKYTCOCGIPQSGSLSLTLC 839  
837 LCFGDMENKLPFAVQDGLLRLFVDDLLVTPHLDQAKTFLSTLVHGVPEYGCMLN 896  
840 LCFGDMENKLPFAVQDGLLRLFVDDLLVTPHLDQAKTFLSTLVHGVPEYGCMLN 899  
897 VYNPFEVPTLGGAPQLPAHCLFPWCGLLDTQLEVPDYSYGAOTSIKSLTFQSV 956  
900 VYNPFEVPTLGGAPQLPAHCLFPWCGLLDTQLEVPDYSYGAOTSIKSLTFQSV 959  
957 PKAGTKMKNLLSVLRKCHGLFLDLQNSLQVNCINIKIFLQAYRPHACVQLPFDQ 1016  
960 PKAGTKMKNLLSVLRKCHGLFLDLQNSLQVNCINIKIFLQAYRPHACVQLPFDQ 1019  
1017 RYRKNLTPFLGIIISQASCCVAILKVNPGMTLK---ASGSFPEAAHMLCVQAFLLK 1073  
1020 HYRKNAPAFLLSIISQASCCVAILKVNPGMTLK---ASGSFPEAAHMLCVQAFLLK 1079  
1074 AHSVYKCLGLGLPATAOKLCKLPEATMTILKAAADPALSTDFQILD 1122  
1080 GHSVYKCLGLGLPATAOKLCKLPEATMTILKAAADPALSTDFQILD 1128

## RESULT 2

3-09-990-080-2  
Sequence 2, Application US/09990080  
Patent No. US20020102686A1  
GENERAL INFORMATION:  
APPLICANT: Morin, Gregg B.  
APPLICANT: Genen Corporation

;; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants  
;; FILE REFERENCE: 018/258C  
;; CURRENT APPLICATION NUMBER: US/09/990,080  
;; PRIOR FILING DATE: 1998-08-03  
;; PRIOR APPLICATION NUMBER: US 09/052,864  
;; PRIOR FILING DATE: 1998-03-31  
;; NUMBER OF SEQ ID NOS: 21  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 2  
;; LENGTH: 1132  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-990-080-2

Query Match 59.4%; Score 3505; DB 10; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316; Indels 52; Gaps 13;  
Matches 719; Conservative 122; Mismatches 260;

Qy 1 MTRAPRCAPVRSLLRSRYREVMPLATFVRRLLGPEGRRLLVQPDGPKIYRTLVLAQCLVCMHW 60  
Db 1 MPEAPRCAPVRSLLRSRYREVMPLATFVRRLLGPEGRRLLVQPDGPKIYRTLVLAQCLVCMHW 60  
Qy 61 GSOPPPADLSFQVSSLSKELVARVQRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
Db 61 DASPPPAAPSFQVSSLSKELVARVQRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
Qy 121 SYLNTVITLTVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
Db 121 SYLNTVITLTVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
Qy 181 TTDIWPVSASYPRTVRVGNFNNLPLQOIKSSSQEAPKPLALPSRGTCKHLSLTS 240  
Db 181 ATQARPPPHAS-GPERRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
Qy 241 VPSAKKARCVFVREBEGP-----HRQVLPDSGKSW-VSPSPARSPEVPTAEKDLSSK 292  
Db 232 LPLPKPRRGAPEBERTPVGGSWAHFGRTRGSDRGFCVVSFAR-----PAEATSL 286  
Qy 293 GKVSDDLSS-GSVCKCHKPSSTLSLSPQONAPQLPFTIETHLYSRGDOGERNLPSF 350  
Db 287 GALSGRTHSHSPVGRHAGHPSTSRPPRMDTPCPVYVETKGFILYSSGD-KEQLRPSF 345  
Qy 351 LLSNLQPNLTGARRLVEIIFLGSRRPTSGPLCRTHLSRRYQWRPLFOQLLVNHAECY 410  
Db 346 LLSLSPSTGTARRLVEIIFLGSRRPTSGPLCRTHLSRRYQWRPLFOQLLVNHAECY 405  
Qy 411 VRLRHGRCFRFANQOQTDAL-----NTSPHMLDLRLHSSPMQVY 452  
Db 406 GVLLKTHCPLRAA---VTPAAGVCAREKPDGSAAPBEEDTPRRLVQLLRQHSSPMQVY 462  
Qy 453 GFLPACLCVKVSGASLWGTNRHRRFRKQLKFIISLGYKGLSLQELMKMKVEDCHLRS 512  
Db 463 GFYRACLRVLVPLWGSRRHRRFRKQLKFIISLGYKGLSLQELMKMKVEDCHLRS 522  
Qy 513 SPCKDRVPAEAHRLRILATFLWLMQTYVQLLSRSPFYITESTFQKRLFFYRKSVWS 572  
Db 523 SPVGVCPVPAEAHRLRILATFLWLMQTYVQLLSRSPFYITESTFQKRLFFYRKSVWS 582  
Qy 573 KLSIGVGRHLERLRLRELSEVREHQTDLAMPICRLRIPKPNGLRPIVNNYSWGT 632  
Db 583 KLSIGVGRHLERLRLRELSEVREHQTDLAMPICRLRIPKPNGLRPIVNNYSWGT 642  
Qy 633 RALGRKQAOHFTQRLTFSMLNYERTKHPHMGSSVGLNDIYRTWRAVLAVALDQ 692  
Db 643 RTRRERKARERTSRVYALFSVINTERRARPDGLLGSVGLDDIHRARWTFVLRQAQDP 702  
Qy 693 TPRMYFVKADVTGAYDAI PQGLVLAVALDQTPRMVFKADVTGAYDAI PQGLV 752  
Db 703 PPELYFVKADVTGAYDAI PQGLVLAVALDQTPRMVFKADVTGAYDAI PQGLV 761  
Qy 753 QVTTLSDLQPYMGQFLKHLQSDA SALRNSVVIQSIIMNBSLSDFLFLHLSVVK 812  
Db 762 HVSTLTDLQPYMGQFLKHLQSDA SALRNSVVIQSIIMNBSLSDFLFLHLSVVK 819

813 IGRDCTVCOQGIPOGSSSTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 872  
 820 IRGKSVQCOQGIPOGSSSTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 879  
 873 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEPGLTGGAPQLPAHCLFPWCGLLLDTOT 932  
 880 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEPGLTGGAPQLPAHCLFPWCGLLLDTOT 939  
 933 LEVFCDSYGAQTSIKTSITFOSVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASLTFRNGFKAGRMNRKLFGLVLRKCHGLFLDLQVNSLQTVCI 999  
 993 NIYKIFLLQAYRFHACVIOQLPDORVRKULTFELGIISSQASCCVAILKVNPGMTLKAS 1052  
 1000 NIYKIFLLQAYRFHACVIOQLPDORVRKULTFELGIISSQASCCVAILKVNPGMTLKAS 1059  
 1053 GS---FPPAAHWCYQAFLLKLAHSHVYIKLGLGLRLTAQKLLCRKLPBATMTILKAAA 1109  
 1060 GAAGPLPSEAVQWLCQAFLLKLTREHVTYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQTLTD 1122  
 1120 NPALPSDFKTILD 1132

SULT 3  
 -09-749-728B-31  
 Sequence 31, Application US/09749728B  
 Patent No. US20020142457A1  
 GENERAL INFORMATION:  
 APPLICANT: Umezawa, Akihiro  
 APPLICANT: Hata, Jun-ichi  
 APPLICANT: Fukuda, Keiichi  
 APPLICANT: Ogawa, Satoshi  
 APPLICANT: Sakurada, Kazuhiro  
 APPLICANT: Gojo, Satoshi  
 APPLICANT: Yamada, Yoji  
 TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMY  
 FILE REFERENCE: 00766.000043  
 CURRENT APPLICATION NUMBER: US/09/749,728B  
 CURRENT FILING DATE: 2001-09-17  
 PRIOR APPLICATION NUMBER: H11-372826  
 PRIOR FILING DATE: 1999-12-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-01148  
 PRIOR FILING DATE: 2000-02-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-07741  
 PRIOR FILING DATE: 2000-11-02  
 NUMBER OF SEQ ID NOS: 80  
 SOFTWARE: Patentin ver.2.0  
 SEQ ID NO 31  
 LENGTH: 1132  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 -09-749-728B-31

Query Match 59.4%; Score 3505; DB 10; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 1 MTRAPRCFVRSLLRSYREVWPLATFVRRLGPEGRVLVQGDPKIYRTLVAQCLVCMHW 60  
 1 MTRAPRCFVRSLLRSYREVWPLATFVRRLGPEGRVLVQGDPAFRAALVAQCLVCFW 60  
 61 GSOPFPADLSFHQVSSIKELVAIVVQRLCERNVYLAFFGELLNEARGGPMFTSSVR 120  
 61 DARPPPAAPSFRQVSLKELVAIVVQRLCERNVYLAFFGELLNEARGGPMFTSSVR 120  
 121 SYLPNTVETLRSVGAWMLLSVGGDDLVLVLLAHCALVLLVPPSCAVQVCGSLYQICA 180  
 121 SYLPNTVETLRSVGAWMLLSVGGDDLVLVLLAHCALVLLVPPSCAVQVCGSLYQICA 180  
 181 TTIDWFSVSAYSRPTFRVGRNFTNLRFLQKIKSSSSRQEAAPKLPALPSRGTKRHLSTSTS 240

RESULT 4  
 US-09-843-676-225  
 ; Sequence 225, Application US/09843676  
 ; Patent No. US20020164786A1  
 ; GENERAL INFORMATION:

181 ATQARPPPHAS-GPRRRLUG-----CERAWNSHVRAGVPLGLPAGARRRGGSASRS 231  
 241 VPSAKKARCYVPRVEGCP-----HROVLPSPGKSM-VSPSPARSPVEYPTAKOLSSX 292  
 232 LPLPRPRRGAAPERTFVQGGSWAHPRGRTRGSDRGFCVSPAR-----PAEATSLE 286  
 293 GKVSLSLS-GSVCKKXPSSTSLLSPRQNAFQLRP-FIETRHFLYSRGGGOERLNPSP 350  
 287 GALSSTRSHSVGRQHAGPPSTSRPPRPMDTCCPPVYATKHFLLYSYSGD-KEQURPSP 345  
 351 LLSNLQPNLTGARRLVEIIFLGSRPSTSGPLCRTHLSRRYQWRPPLFQOLLVNHAEQY 410  
 346 LLSLRPSLTGARRLVEIIFLGSRPMPGTPRRLPRLPQRYQWRPPLFLELLGHAQCPY 405  
 411 VLLRSHCRFRANQOVTDAL-----NTSPHMLDLRLHSSPMQVY 452  
 406 GVLLKTHCPLRAA---VTPAAGVCAREKPOGSAAPBEEDTDPRRLVQLLRQHSSPMQVY 462  
 453 GFLRACLCKVYSASLWGTNRHNRFFKNLKKFISLGRYKLSLQELMWMKMKVEDCHWLS 512  
 463 GFVRACLRLVPPGLWGRNRHNRFFLNTKKFISLGRYKLSLQELMWMKMKVEDCHWLS 522  
 513 SPKDRVPAABERLPERILATFLPWLMDTYVYVLLRSFFYITESTFOKNLFFYRKSVM 572  
 523 SPGVGCVZAAERLREILAKFLHLMVYVVELRSFFYVTTETFOKNLFFYRKSVM 582  
 573 KLOSIGVRCHLRLVRLRELSOEVRHDDTWLAMPICRLRFIPKENGRLPIVNSVSGT 632  
 583 KLOSIGIRQHLKRVQLRELSAEVRQREARPAULTSLRLRFIPKPDGLRFVINDYVYGA 642  
 633 RALGRKQAOHTORLKLTFSLMNYERTKPHLMGSSVLGMNDIYRTWRAFLVLRALDQ 692  
 643 RTFREKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRAMRTFLVLRADP 702  
 693 TPRMYVADVTGAYDAIPQKLVVVANMIRHSESTYCIQYAVVTRDSCQVHKSPFR 752  
 703 PPELYFVKVDVTGAYDTIPQRLTEVIAIHK-PONTYCVRRYAVVQAAHGHVYKAFKS 761  
 753 QVTTLSLDQVYMGQKHLQSDASALRNSVVIQOSISMNESSSDFLHFLRHSHVWK 812  
 762 HVTSLTDLPYMRQFVAHQET--SPLEDAVVIQOSSSLNEASSGLFDVFLRFMCHAVR 819  
 813 IGRDCTVCOQGIPOGSSSTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 872  
 820 IRGKSVQCOQGIPOGSSSTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 879  
 873 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEPGLTGGAPQLPAHCLFPWCGLLLDTOT 932  
 880 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEPGLTGGAPQLPAHCLFPWCGLLLDTOT 939  
 933 LEVFCDSYGAQTSIKTSITFOSVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASLTFRNGFKAGRMNRKLFGLVLRKCHGLFLDLQVNSLQTVCI 999  
 993 NIYKIFLLQAYRFHACVIOQLPDORVRKULTFELGIISSQASCCVAILKVNPGMTLKAS 1052  
 1000 NIYKIFLLQAYRFHACVIOQLPDORVRKULTFELGIISSQASCCVAILKVNPGMTLKAS 1059  
 1053 GS---FPPAAHWCYQAFLLKLAHSHVYIKLGLGLRLTAQKLLCRKLPBATMTILKAAA 1109  
 1060 GAAGPLPSEAVQWLCQAFLLKLTREHVTYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQTLTD 1122  
 1120 NPALPSDFKTILD 1132

APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Harley, Calvin  
 Andrews, William H.

TITLE OF INVENTION: No. US20020164786A1el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/843,676

FILING DATE: 26-Apr-2001

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 0153899-0029300US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 225:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 225:

S-09-843-676-225

Query Match 59.4%; Score 3505; DB 10; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.66-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

Y 1 MTRAPRCAPVSLRSRYREYVPLATFVRRLGPEGRRLVQGPDKYRTILVAOCLVCMHW 60  
 b 1 MTRAPRCAPVSLRSRYREYVPLATFVRRLGPEGRRLVQGPDKYRTILVAOCLVCMHW 60

Y 61 SQPPADLSHQVSSLSKELVARVQRLCERNENVLAFQFELLNARGGPPMAFTSSVR 120  
 b 61 DARPPPAAPSPRQVSCLELVARVQLRCERGANVLAFGFALLDARGGPPPEAFTTSVR 120

Y 121 SYLPTNTVETLRVSGAWMLLSRYGDDLLVYLLAHACALYLLVPPSCAYQVCGSPLYQICA 180  
 b 121 SYLPTNTVETLRVSGAWMLLSRYGDDLLVYLLAHACALYLLVPPSCAYQVCGSPLYQICA 180

Y 181 TTDIWPVSASVYRTPVGRNFTNLRFLQKISSRQAPKALPSRGTKRHLSTSTS 240  
 b 181 ATQARPPPHAS-GRPRRLG-----CERAWNHSVREAGVPLCLPAPGARRRGGSASRS 231

Y 241 VPSAKKARCYVPVPRVEEP-----HRQVLPTSGKSW-VPSPARSPVPTAEKLSK 292  
 b 241 VPSAKKARCYVPVPRVEEP-----HRQVLPTSGKSW-VPSPARSPVPTAEKLSK 292

Y 232 LPLPKRPRGAPEPERTPVGGGWAHPGRTGRGFCVVSFAR-----PAEATSLE 286  
 b 232 LPLPKRPRGAPEPERTPVGGGWAHPGRTGRGFCVVSFAR-----PAEATSLE 286

## RESULT 5

US-09-953-052-2

; Sequence 2, Application US/09953052

; Patent No. US20020173476A1

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; Lingner, Joachim

; Nakamura, Toru

; Chapman, Karen B.

; Morin, Gregg B.

; Harley, Calvin B.

QY 293 GKVSLSLS-GSVCKHKPSSTSLSPRQNAFOLRP-FIETRFHLYSRGGOERLNPSP 350  
 DB 287 GALSOTRSHSPSVGQHGAGPPSTSRPPRWDTPCPPIYATKTHLYSSGD-KEQLRPSF 345  
 QY 351 LLSNLQPNLTGARRLVEIIFLGSRPRTSGPLCRTHLSRRYQWRPLFQOOLLVNHAEQY 410  
 DB 346 LLSLSRPSITGARRLVEIIFLGSRPWPGTTPRRLPRLPQRYQWRPLFLELLGNHACPY 405  
 QY 411 VLLRSHCRFTANQOVTDAL-----NTSPHMLDLRLHSSPMQVY 452  
 DB 406 GVLKTHCPCLRAA---VTPAAGVCAREKPGSVAAPEEEDTPRRLVQLLRQHSPPQVY 462  
 QY 453 GFLRACLCKVVSASLWGTNRHNERFFKNLKFISLGYKGLSLOELMWKMKVEDCHWLRS 512  
 DB 463 GFVRACLRLVPPGLWGSRRHNERFLNKKFISLGKXAKLSLOELTWKMSVRDCAWLR 522  
 QY 513 SPKORVPAAEHRLRERILATFLWMDTYVQLRSPFFYITESTFOKNRLFYFKKSVMS 572  
 DB 523 SPQVGCVPAAEHRLREEILAKFLHWSVYVVELLSFFYVTTTFOKNRLFYFKKSVMS 582  
 QY 573 KLSIGVROHLERVLRELSEVSHODTWLAMPICRLREIPKPNGLRPTVNMYSMGT 632  
 DB 583 KLSIGIROHLERVLRELSEAEVQHREARPAALLTSELRFIPKPDGLRPIVNDYVUGA 642  
 QY 633 RALGRKQAOHFTQRLKTLFSLMNTYRKHPLMGSSVLGNDIYRTWRAPVLRYRALDQ 692  
 DB 643 RTFRREKRAERLTSRVKALFSLVINYERARRPCLLGASVLGLDDIHRARWTFVLVRADP 702  
 QY 693 TRMPEVKADVTGAYDAIPQKLVWVANMIRHSESTYCIROYAVVRDSDQGVHKSFR 752  
 DB 703 PPELYFVKVDVTGADTTPQDLTEVIASIIK-PQNTYCVRYAVVQKAAHGVYKAFKS 761  
 QY 753 QVTTLSDLQPYMGQFLKHLQSDASALRNSVVI EOSISMNESSSLEDFFLHFLRHSVVK 812  
 DB 762 HVSTLTDLPYMEQFVAHLQET--SPLRDVAVIEQSSSLNEASSGLEFDFVLFEMCHAVR 819  
 QY 813 IGDRCTYQCGIPQSSSLTLLCSLCFGDMENKLPFAEVORDGLLLRFVDDFLLYTHLDQ 872  
 DB 820 IRGKSYVQCGIPQSSIJLTLCSLCYGMENKUFAGIRRDGLLLRLVDDFLLYTH 879  
 QY 873 AKTFSLTLVHGVPYEGCMINLQNTVYVFPFPGTLGGAAPYQLPAHLFPWCGLLDTQT 932  
 DB 880 AKTFSLTLVHGVPYEGCVNLRKTVNPFVEALGGTAFVQMPAHGLFPWGLLDTRT 939  
 QY 933 LEVFCDSYGAOTSINTSLTFSQVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLTQVCI 992  
 DB 940 LEVQSDYSYARTSIRASLTFNRGFKAGNMRKLFGLVRLKCHSLFLDLQVNSLTQVCI 999  
 QY 993 NIYKIFLLQAYRPHACVIOQLPFDQVRKRLTFPLGIISSQASCCVAILKVNPGMTLKAS 1052  
 DB 1000 NIYKILLQAYRPHACVIOQLPFDQVRKRLTFPLGIISSQASCCVAILKVNPGMTLKAS 1059  
 QY 1053 GS---FPPEAAHWLCYQAFLLKLAHSVIYKCLIGPLRTAOKLCKRKLPEATMTILKAAA 1109  
 DB 1060 GAAGPLPSEAVQWLCQAFLKLLTSHRVYTPVPLGSLRTAQQLSRKLPGTTLLALEAAA 1119  
 QY 1110 DPALSTDFQITLD 1122  
 DB 1120 NPALPSDFKITLD 1132

Andrews, William H.  
 TITLE OF INVENTION: Antisense Compositions for Detecting and  
 Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Townsend and Crew LLP

APPLICANT NUMBER: US/09/953,052

FILING DATE: 14-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/052,919

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/974,549

FILING DATE: 19-NOV-1997

APPLICATION NUMBER: US 08/974,584

FILING DATE: 19-NOV-1997

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Parent, Annette S.

REGISTRATION NUMBER: 42,058

REFERENCE/DOCKET NUMBER: 015389-00360005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

-09-953-052-2

Query Match

Best Local Similarity

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCAPVRLSRVREVMPLATFVRLGPEGRRLVQDPDKIYRTILVAQCLVCMHW 60

1 MTRAPRCAPVRLSRVREVMPLATFVRLGPEGRRLVQDPDKIYRTILVAQCLVCMHW 60

61 GSQPPADLSFQVSSKLVARVVRICERNVLAFGPELLNEARGGPPMAFTSSVR 120

61 DARPPAPSPFQVSSKLVARVVRICERNVLAFGPELLNEARGGPPMAFTSSVR 120

121 SYLNTVETLRVSGAWMLLSRVGDDLLIVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
 121 SYLNTVETLRVSGAWMLLSRVGDDLLIVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
 181 TTDIWPVSASRYRTPVGRNFTNLRFLQCIKSSSQEAPKPLALPSRGTKRHLSTLST 240  
 181 ATQARPPPHAS-GRRRRLG-----CERAWNHVSREAGVPLGLPAPGARRRGGASRS 231  
 241 VPSAKKARCYVPVREVEGP-----HRQVLPTPSKSM-VPSPARSPVPTAEKOLSK 292  
 232 LPLPKRPRGAAPFERTPVGQSWAHPGTRGSDRGFCVWSFAR-----PAEEATSL 286  
 293 GKVSLSLS-GSVCKHKPSSTLSLSPRONAQQLRP-FIETRHFLYSRGDGOERLNPSF 350  
 287 GALSGRHSHPSVGRQHAGPPSTSRPPRPWDTPCPVVAETKHFLYSSGD-KEQLRP 345  
 351 LLSNLQPNLTGARELVEIIFLGRPRPTSGPLCTHRLSRRYQMRFPOQLLVNHAECQ 410  
 346 LLSLRPLSLTGARLVEIIFLGRPRPTSGPLCTHRLSRRYQMRFPOQLLVNHAECQ 405  
 411 VLLRSHCRFTANQVTDAL-----NTSPHLMDDLRLHSSPWQY 452  
 406 GVLKTHCHFLRAA---VTPAGVCAREKPGSVAAPEDDTPRRLVQLLRQHSWPQY 462  
 453 GFLRACLCVVVASLWGRHNRFFKNLKKFISLGKYGKLSLQELMWKVKVEDCHWLS 512  
 463 GFVACLRLRVPCLWGSRNERRFLNTKKFISLGKHAKLSQLTLMKSVRDCAWLAR 522  
 513 SPGKDRVPAAEHRLRERILATFLWMDTVVQLLSFFYIIESTFKNRLPFYRSVNS 572  
 523 SPGVGCVPAAEHRLRERILAKFLWLSVVVVELLSFFYVTTTQKRLPFYRSVNS 582  
 573 KLSIGTVRQHLRVRRLRELSOEVRHHDQTLWAMPICRLRFIPKPNGLRPIVNMYSMT 632  
 583 KLSIGTVRQHLRVRRLRELSOEVRHHDQTLWAMPICRLRFIPKPNGLRPIVNMYSMT 642  
 633 RALGRKQAOHFTORLKTFLSMLNRYTKPHLMGSLVGLGMDIYRTWRAFLVRALDQ 692  
 643 RTFRREKAERLTSRVKALFSLNRYERARPPGLGASVLGLDDIHEARFTFLVRVAQDP 702  
 693 TPRVYFKADVGTGADAIPOGKLVVVVNMIRHSBETCYCIQAVVVRDSQGOVHKSFR 752  
 703 PPVYFKADVGTGADAIPOGKLVVVVNMIRHSBETCYCIQAVVVRDSQGOVHKSFR 761  
 753 QVTLSDLPYMGQFLKHLQSDASALRNSVIEQSI SMNNESSLSLDFFLHPLRHSVYK 812  
 762 HVSTLTDLPYMGQFLKHLQSDASALRNSVIEQSI SMNNESSLSLDFFLHPLRHSVYK 819  
 813 IGRCYTCQCGIPGSSSLTLLCSLCEGDMENKLFPAEQVDGLLRFVDDFLVTPHLDQ 872  
 820 IRGKYVQCGIPGSSSLTLLCSLCEGDMENKLFPAEQVDGLLRFVDDFLVTPHLDQ 879  
 873 AKTFLSLTVHGVPEYGCWMLQKTVNFPVEPGTGLGAAAPYQLPAHCLFPWGLLDTQ 932  
 880 AKTFLSLTVHGVPEYGCWMLQKTVNFPVEPGTGLGAAAPYQLPAHCLFPWGLLDTQ 939  
 933 LEVPCDYSVTAQTSIKTSLTQSVFVAGKTMKNSLVRLKCHGLFILDQVNSLQTVCI 992  
 940 LEVQDYSVTAQTSIKTSLTQSVFVAGKTMKNSLVRLKCHGLFILDQVNSLQTVCI 999  
 993 NIYKIFLLQAVRFHACVQLPFPDQVRKNTLTFPLGIISSOASCCVAILKVKQPGMTLKA 1052  
 1000 NIYKIFLLQAVRFHACVQLPFPDQVRKNTLTFPLGIISSOASCCVAILKVKQPGMTLKA 1059  
 1053 GS---FPPEAHMICYQAFLKLAASHVYKLLGGLPRTAQKLLCKLPKLPATMTILKAAA 1109  
 1060 GAAGFLPSEAVOMLCHQAFLLKLTTRHRTVYVPLGLSLRTAQKLLCKLPKLPATMTILKAAA 1119  
 1110 DPALSTDQFILD 1122  
 1120 NPALSDQFILD 1132

SULT 6  
-10-295-681-57  
Sequence 57, Application US/10295681  
Publication No. US20030166270A1  
GENERAL INFORMATION:  
APPLICANT: E. Premkumar Reddy  
APPLICANT: Sushil G. Kane  
APPLICANT: Richard V. Mettus  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REVERSIBLY  
INDUCING CONTINUAL GROWTH IN NORMAL CELLS  
FILE REFERENCE: 6056-307  
CURRENT APPLICATION NUMBER: US/10/295,681  
CURRENT FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: US 60/334,760  
PRIOR FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 57  
LENGTH: 1132  
TYPE: PRT  
ORGANISM: Homo sapiens  
-10-295-681-57  
Query Match 59.4%; Score 3505; DB 12; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
/ 1 MTRAPRCNVRSLRSRYREVWPLATFVRLGPEGRLVQPGDKPIRYTLVAQCLVCMHW 60  
/ 1 MPRAPRCNVRSLRSRYREVWPLATFVRLGPEGRLVQPGDKPIRYTLVAQCLVCMHW 60  
/ 61 GSQPPADLSHFQYSSLSKELVARVQVRLCERNERNVLAFCGELLNARGGPMATTSVR 120  
/ 51 DARPPAPSPRQVSCUKELVARVQLCERGAKNVLAFCGALLDARGGPEATTSVR 120  
/ 121 SYLNTVITLVRSGAMWLLSRVDDLLVLAHALYLLVPPSCAYOVQCSPIYQICA 180  
/ 121 SYLNTVITDARGSGAMWLLSRVDDLLVLAHALYLLVPPSCAYOVQCSPIYQICA 180  
/ 181 TTDIPSVASRYVTRPVGRNFTNLRLQIKSSRCQEAPEKPLALPSRGTKEHLSITS 240  
/ 181 ATQARPPHAS-GPRRLG-----CERAWHVSREAGVFLGLPAPARRGGASAS 231  
/ 241 VPSAKKARCPVPVREBP-----HRQVLTPTSGKSM-VPSPARSPVPVTAEXDLSK 292  
/ 232 LPLPKRPRGAAPERTPTVQSGSWAHFGRTRGSDRGFCVWSPAR-----PAEATISLE 286  
/ 293 GKVSLSLS-GSVCKKPSSTLSLSPRQNAQLRP-FIETHLYSRGDCQERLNPSF 350  
/ 287 GALSGRHSFVSQHHAGPSTSRPRPDWTPCPVYAEIKHFLYSSGD-KEQLRPSF 345  
/ 351 LLSNLQNLGTGARRLVETIFLGSRPRTSGPLCRTHRLSRYWQVRPLFOQLLVNHAECQY 410  
/ 346 LLSLRPSLTGARRLVETIFLGSRPWPGTPRRLPLRQYQWVRPLFLELLGNHAQCPY 405  
/ 411 VALLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
/ 406 GYLLKTHCPLEAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLVQLLRQSSPWQY 462  
/ 453 GFLRACLCKVYSASLWGRHNERFFKNKKFISLGYKSLQELMWKMKVDDCHWLS 512  
/ 463 GFLRACLRLVPPGLWSRHRNERFFLNTKFFISLGHAKLSLQELTWMSVDDCANLRR 522  
/ 513 SPGRDVRPAEHRLEBRLATFLFWMNDTVVQLRSFFVITESTQKRLFFYRKSVM 572  
/ 523 SPGVGVPAEHRLEBRLATFLFWMNDTVVQLRSFFVITESTQKRLFFYRKSVM 582  
/ 573 KLOSIGVCHLRLVRLSREOEVHRHODTWLAMPICLRFIPKNGRLPVMNWSMGT 632  
/ 583 KLOSIGVCHLRLVRLSREOEVHRHODTWLAMPICLRFIPKNGRLPVMNWSMGT 642  
/ 633 RALGRKQAOFTORLKLFLGMLNRYETKHPHLMGSSVLGMNDIYRTWRAFLVRLALDQ 692

643 RTFREKRAERLTSRVKALFSLNVERARRPCLLGASVLGLDDIHRARWTFVLVRADQP 702  
693 TPRMYFVKADVTGAYDAIPQKLVVWVANMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
703 PPELYFVKVDVTGAYDTIPQRLTEVIASIIK-PQNTYCVRRYAVVQKAAGHYRKAFK 761  
753 QVTTLSLDLOPMYQGLKHLQSDASALENSVVISQISMNESSSLDFDLFLRHSVVK 812  
762 HVTSTLDDQPMYRQFVAHLQET--SPLDAAVVIQSSSLNEASSGLDFDLRFMCHAVR 819  
813 IGDRCYTCQGIPOGSSSLTLLCSLCFCGDMENKLFQVQDRLLRVDDPFLVTPHLQ 872  
820 IRGKSVQCGIPQGSILSTLLCSLCYCGDMENKLFAGIRRDGLLRVDDPFLVTPHL 879  
873 AKTFSTLVHGVPEYGCMMINLOKTVNPEVPGTLLGGAAPYQLPAHCLFPMWGLLDTOT 932  
880 AKTFSTLVHGVPEYGCMMINLOKTVNPEVPGTLLGGAAPYQLPAHCLFPMWGLLDTOT 939  
933 LEVFCYSGYAOTSIKTSLTFQSVFVFKAGTKMENKLLSVLRKCHGLEFLDQVNSLOTVCI 992  
940 LEVQSDYSSYARTSIRASITFNRGFKAGRNWRKLFGLVRLKCHSLFLLDQVNSLOTVCT 999  
993 NYKIFLQAYRFHACVQLQPPQDQVRKNLTFPIGIISSQASCCYAILKVNKPMOTLKAS 1052  
1000 NYKIFLQAYRFHACVQLQPPQDQVRKNLTFPIGIISSQASCCYAILKVNKPMOTLKAS 1059  
1053 GS---PPPEAAHLCVQAFLLKLAHSAHVIYKCLLPRLTAQKLCRKLPEATMTILKAAA 1109  
1060 GAAGPLPSEAVONLCHQAFLLKLTHTVYVYVPLGLSLTAQTLQSLRKLPGTTLTALEAAA 1119  
1110 DPALSTDFOTILD 1122  
1120 NPALPSDFKTILD 1132

RESULT 7  
US-10-053-758-225  
; Sequence 225, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Linsinger, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION DATA:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
-10-053-758-225

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCRAVRSLLRSRYREVPLATFVRRLLQPEGRRLVQPDGPKIYRTLVAAQCLVCMHW 60  
1 MPAPRCRAVRSLLRSRYREVPLATFVRRLLQPEGRRLVQPDGPKIYRTLVAAQCLVCPW 60  
61 GSOPPPADLSFQVSSLSKELVARVVRQRCERNRNVLAFGELLNEARGGPPMAFTSSVR 120  
61 DARPPAPAFSPFQVSCLELVARVLQRCERAKNVLAFFGALLDGGGPEAFITTSVR 120  
121 SYLNPVTIELRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLVYOICA 180  
121 SYLNPVTIELRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLVYOICA 180  
181 TTDIMPVSAVSRPVRGNTNRLFLQIKSSRQEAQPKLALPSGTRKHLSLSTGS 240  
181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGAPGARRRGGSASRS 231  
241 VPSAKKARCVPRVVEEGP-----HRQVLTSPGKSM-VPSPARSPVPTAEKDLSSK 292  
232 LPLPKPRGRGAPEPRTVVGSGWAHPGRTGSDRGFCVVSAP-----PAAEATSL 286  
293 GKVDLSLS-GSYCCCKHPSSTLSPPRQNAFLRP-FIETRIPLXSRGDGQERLNPSP 350  
287 GALSGRHSHPSVGRHAGPSTSRPPRPMDTPCPVVAETKHFLYSSGD-KEQLRPSF 345  
351 LLSNLOPNTIGARLVEIIFLSRPTSGPLCRHLSRRYQWRPPLFQQLLVNHAECQY 410  
346 LLSLSLPSLTGARRLVETIFLGSRPMPGTPRRLPRLPQRYQWRPPLFLELLGHAQCPY 405  
411 VRLLRSHCRFRFTANOOVTDAL-----NTSPHMLDLLRHSSPMQVY 452  
406 GVLLKTHCPLRAA---VTPAAGCAKEKPGQSVAAPESEDTPRLVQLLRHSSPMQVY 462  
453 GFURACLCKVVASLWGTTHNRRRFFKNLKKPISLGKYGKLSLOELMMKMKVEDCHMLRS 512  
463 GFVRACLRLVPPGLWGSNRRRLRNLTKEFISLGKEAKLSLOELTWKMSVRDCAMLR 522  
513 SPCKDVPAAEHLRLERILATLFWLMDTVVQLARSFFYTESTFOKNRFFFRKSVWS 572  
523 SPVGCVPAEHLRLERILAKFLHMLSVYVVELLRSFFYVETTFQKNRFFFRKSVWS 582  
573 KLSIGVROHLRLERILRSQBEVRHODTWLAMPICRLRIFPKPNGLRPIVNMYSMTG 632  
583 KLSIGIRQHLKRVQLRELSEAVRQREARPAALLTSRLRIFPKPDGLRPIVNMVYVGA 642  
633 RALGREKQAHFORLKTFLSMNTERYKHPHLMGSSVLGWNDDIYRTWARVLRALDQ 692  
643 RTRPREKRAERTSRVKALFSVLNTERARRPGLLGASVLGDDIHRAWRTFVLRVRAQDP 702  
693 TRPMYFVKADVTGAYDAIPQGLKVEVAVNMRHSESTYCIQYAVVRDSDGQVHKSFRR 752

Db 703 PPELYFVKVDVTGAYDITIPODRLTEVTASIIK-PQNTYCVRRYAVVQKAAHGHVRKAFKS 761  
QY 753 QVTTLSDLQPYMGOFKLKHLQSDASALRNSVVEIQSISMNNESSSLDFDFLHFLRHSVVK 812  
Db 762 HVSITLTDLQPYMRQFVAHQET--SPURDAVVEIQSSSLNEASSGLFDVFLRFMCHHAVR 819  
QY 813 IGDRCYTQCOGIPQSSSLTLLCSLCFGDMENKLFARVQDQDGLLLRFPVDDFLVTPHLDQ 872  
Db 820 IRGKSYVQCOGIPQSSSLTLLCSLCYGDMDENKLFAGIRDDGLLLRLVDVDFLLVTPHLTH 879  
QY 873 AKTFLSTLVHGVPEYGCMINLOKTVNFPVEPGTGLGAAPYQLPALHCLFPWCGLLLDTQT 932  
Db 880 AKTFLRTLVRGVPEYGCVINLRKTVNFPVEDEALGCTAFVQMPAHGLFPWCGLLLDTRT 939  
QY 933 LEVFCDSYGYACTSIKTSITPQSVFKAGKTMRNKLSVLRLKHGHLFLLDQVNSLQTVCI 992  
Db 940 LEVQSDYSYARTSIRASLTENRGFKAGRNNRKLFGVLRKCHSLFLLDQVNSLQTVCT 999  
QY 993 NIYKIFLLQAVRFHACVIOLPFDORVRKNLTFFLGIISSOASCYAILKVKNPGMTIKAS 1052  
Db 1000 NIYKILLQAVRFHACVIOLPFHQQVWKNFTFFLRVISDTSASLCYSILKAKNAGMSLGAK 1059  
QY 1053 GS---FPPEAAHFLCYQAFLLKLAHSAVYIKCLGLPLRTAKLLCRKLPEATMTILKAAA 1109  
Db 1060 GAAGPLPSEAVOMLCHQAFLLKLTTRHRTVTVPLIGLSLRTAQTQLSRKLPGTTLTALAAAA 1119  
QY 1110 DPALSTDFQTILD 1122  
Db 1120 NPALPSPDFKTILD 1132

RESULT 8  
US-10-208-243-2  
; Sequence 2, Application US/10208243  
; Publication No. US20030044394A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaeta, Federico C.A.  
; APPLICANT: Genon Corporation  
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune  
; TITLE OF INVENTION: Response to a Telomerase Antigen  
; FILE REFERENCE: 015389-003500PC  
; CURRENT APPLICATION NUMBER: US/10/208,243  
; PRIOR FILING DATE: 2002-07-30  
; PRIOR APPLICATION NUMBER: US/09/675,321  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/112,006  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898  
; PRIOR FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1132  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-208-243-2

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCRAVRSLLRSRYREVPLATFVRRLLQPEGRRLVQPDGPKIYRTLVAAQCLVCMHW 60  
Db 1 MPAPRCRAVRSLLRSRYREVPLATFVRRLLQPEGRRLVQPDGPKIYRTLVAAQCLVCPW 60  
QY 61 GSOPPPADLSFQVSSLSKELVARVVRQRCERNRNVLAFGELLNEARGGPPMAFTSSVR 120  
Db 61 DARPPAPAFSPFQVSCLELVARVLQRCERAKNVLAFFGALLDGGGPEAFITTSVR 120  
QY 121 SYLNPVTIELRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLVYOICA 180  
Db 121 SYLNPVTIELRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLVYOICA 180

181 TTDIWPVSASVPRPVRNFTNLRFLQKIKSSROEAKPLALPSRGTGKHLSTSTS 240  
181 ATQARPPPHAS-GPRLRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
241 VPSAKKACVFPVREVEGP-----HRQVLPTEPSGKSM-VPSPARSPVPTAEKDLSSK 292  
232 LPLPKRPRGAPEPRTFVQGGWNAHGRTRGSDRGFCVVSAPAR-----PAEZATSL 286  
293 GKVSLSLS-GSVCKKXPSSTLSLSPRQNAFOLRP-FIETHFYSGDQGERLNPSF 350  
287 GALSCTRHSHPVSGHQHAGPSTSRPRPMDTCCPPVYAEATKHFLYSSGD-KEQLRPSF 345  
351 LLSNLQPNLTGABRLVETIFGSPRSTSGPLCRTHLSRRAVQWMPPLFOQLLVNHAECQY 410  
346 LLSLRPSLTGABRLVETIFGSPRSTSGPLCRTHLSRRAVQWMPPLFELLNHACQPY 405  
411 VLLSHCRFTANQVTDAL-----NTSPPHMLDLRLHSSPWQVY 452  
406 GVLLKTHCPRLAA---VTPAAGVCAREKPGQSVAAPEEDTDPRRLVQLLRQHSSPWQVY 462  
453 GLRACLCKVYSASLWGTNRHNRFFKNLKFISLGYKGLSLOELMKWKVEDCHLRS 512  
463 GVRACLRLVPPGLWGRHNRFLNRKFKFISLGHAKLSLOELTKWNSVRDCAWLR 522  
513 SPGKDRVPAAEHRLERILATLFLWMDTYVQLRSTFFVITESTFOKNLFFYRKSVM 572  
523 SPGVCVPAAEHRLERILATLFLWMDTYVQLRSTFFVITESTFOKNLFFYRKSVM 582  
573 KLOSTGVQHLERVLRELSEVVRHODTWLAMPICLRFPKPNGLRPIVNMYSMGT 632  
583 KLOSTGIRQLKRVQLRELSAEVRQREARPAULTSLRFPKPNGLRPIVNMYSMGT 642  
633 RALGRKQAOHTORLKTILFSLMAYERTKPHLMGSSVLGMNDIYRTWRAVLRVRLDQ 692  
643 RTREREKAEALTSKALFSLVNYERARSGLLGASVLSDDIHRARWTFVLRAQDP 702  
693 TPRMTFVADVTGADAIPOKLVVEVNMTRHSESTYCIQAVVRDSDQGVHKSFR 752  
703 PPELVFVKVDVTGADITFQDRLTEVIAIILK-PQNTYCVRRYAVVQAAHGHVRKAFKS 761  
753 QVTTLSLDLPVYMGQFLKHLQSDASALRNSVVIQSI SMNSSSLPDDFLFLHLSVVK 812  
762 HVSTLTDLPVYMGQFLKHLQSDASALRNSVVIQSI SMNSSSLPDDFLFLHLSVVK 819  
813 IGRCYTCQGIPOGSSSTLSCISCFDGMENKLF AEVQDGLLRFLVDDFLVTPHLDQ 872  
820 IRGKSVQCOGIPGSSSTLSCISCFDGMENKLF AEVQDGLLRFLVDDFLVTPHLDQ 879  
873 AKTELSTLVHGYPEYGCMLNLOKTVNFPVPGTILGGAAPVQLPAHCLFPWCGLLDQT 932  
880 AKTELSTLVHGYPEYGCMLNLOKTVNFPVPGTILGGAAPVQLPAHCLFPWCGLLDQT 939  
933 LEVFCDSYGAOTSITSLTQSFVFKAGTKMRNKLKSLVRLKCHGLFLDLQVNSLTQVCI 992  
940 LEVQSDYSYARTSIRASITFNRGFKAGENMRKLFGLVRLKCHSLFLDLQVNSLTQVCT 999  
993 NIYKIFLLQAYFPHACVQLPDPQVRQLTFFLGISSQASCCYAILKVNQWMLKAS 1052  
1000 NIYKIFLLQAYFPHACVQLPDPQVRQLTFFLGISSQASCCYAILKVNQWMLKAS 1059  
1053 GS---PPPEAAHLCYQAPFLKLAHNSVYKCLGPLRTAQKLCRLKPEATWTILKAAA 1109  
1060 GAAGPLPSEAVQWLCQAPFLKLAHNSVYKCLGPLRTAQKLCRLKPEATWTILKAAA 1119  
1110 DPALSTDPOTILD 1122  
1120 NPALPSDFKILD 1132

## RESULT 9

US-10-054-295-225

Sequence 225, Application US/10054295

Publication No. US20030044953A1

GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Hartley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030044953A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,295  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
US-10-054-295-225

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPCPAVRSILRSRYREVWPLATFVRRLGPEGRLVQPGDKIYRLVLAQCLVCMHW 60  
Db 1 MPRAPCPAVRSILRSRYREVWPLATFVRRLGPEGRLVQPGDKIYRLVLAQCLVCMHW 60  
QY 61 GSQPPADLGFHQVSSLKELVARVORLCERNERNVLAFFGELLNEARGGPPMATSSVR 120  
Db 61 DARPPAPSPRQVSCIKELVARVORLCERNERNVLAFFGELLNEARGGPPMATSSVR 120  
QY 121 SYLNTVETLRVSGANWMLLSRVGDDLVLLAHALYLLVPPSCAYOVCGSPYQICA 180  
Db 121 SYLNTVETLRVSGANWMLLSRVGDDLVLLAHALYLLVPPSCAYOVCGSPYQICA 180  
QY 181 TTDIWPVSASVPRPVRNFTNLRFLQKIKSSROEAKPLALPSRGTGKHLSTSTS 240  
Db 181 ATQARPPPHAS-GPRLRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
QY 241 VPSAKKACVFPVREVEGP-----HRQVLPTEPSGKSM-VPSPARSPVPTAEKDLSSK 292



232 LPLPKPRGAAPEBPTPVGGSWAHQRTGRGSDRGFCVVSPAR-----PAEATSLT 286  
 293 GKVDLSLS-GSYCKKPSSTLSPPRQNAQLRP-FIETHFLYSGDQGRNLNPSF 350  
 287 GALSCTRHSHPSVGRHAGAPSTSRPPRPWDTPCPVVAETKHFYSSGD-KEQLRPSF 345  
 351 LLSNLQNTGARRLEIIFLGSRRPTSPCLRTHLSRRYQWEPFLQQLLVNHAECQY 410  
 346 LLSLSRPSGTARRLVETIFLGSRRPMPGTPRRLPRLPQRYQWEPFLLELLGHAQCPY 405  
 411 VRLRSHGCRFTANQOVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
 406 GVLLKTHCLPRAA--VTPAAGVCAREKPGQSVAAPEBEDTDRVLQVLLQHSSPWQVY 462  
 453 GFTBACLCKVVSASLWGRHNRERRFFKNLKFISLQKYGKLSQLBLMKWQKVEDCHWLS 512  
 463 GFVRACLRRLVPPGLWGRHNRERRFLNFKFISLQKHLKLSQLBLTKWMSVRCWLRR 522  
 513 SPQKDRVPAEHRRLREILATFLFWMMDTVVYVOLLRSFFYITESTFQKNRFFFYKSVWS 572  
 523 SPQVGCVPAAEHRRLREILAKFLWLMMSVYVVELLRSFFYITETTFQKNRFFFYKSVWS 582  
 573 KLOSIGVROHLERVLRLSLSQEVRRHQTWAMPICRLRFPKPNGLRPIVNNYSNGT 632  
 583 KLOSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFPKPDGLRPIVNNDYVGA 642  
 633 RALGRKQAOHFTQRLKTLFSLMYERTKPHLMGSSVLGMDIYRTWRAFVLRVLDQ 692  
 643 RTPRERKARELTSRVKALFSLVNERARPPGLGASVLGDDLIHRAWRTFVLVRQDPP 702  
 693 TPRLMYFVADVTGAYDAIPQGLVEYVANMIRHSESTYCIQYAVVRDSQGVHKSFR 752  
 703 PPBLFYKVDVTGAYDTIPQDRLETVIASIK-PQNTYCVRRYAVVQAAHGHVKAFKS 761  
 753 QVTLSDIQPYMGQFLKHLQSDASALRNSVIEQISWNBSSSLFPDFLFLRHSVVK 812  
 762 HVSTLTDLPYMRQFVAHQET--SPLRDAVIEQSSSLNEASSGLFDVFLRFVCHAVR 819  
 813 IGRCYVTCQGIQPGSSSLTLCSCFCGDMENKLFPAEVORQGLLRFVDDLLVTPHLQ 872  
 820 IRKSYVOCQGIQPGSILSTLCSCYCDMENKLFAGIRRDGLLRLVDDLLVTPHL 879  
 873 AKTFLTLVHGPEYGCMLNOKTVNFPVEPGTIGGAAPYQLPACHLFPWCGLLLTQT 932  
 880 AKTFLTLVHGPEYGCMLNOKTVNFPVEDEALGGTAFVQMPAHGLFPWCGLLLTQT 939  
 933 LEVPCDYSYVAQTSITSLTFQSVKAGTKMENKLLSVLRKCHGLFLDLQVNSIQTVCI 992  
 940 LEVQSDYSYVAQTSIRASITFNRGFKAGNMRKLFGLVLRKCHSLFLDLQVNSIQTVCT 999  
 993 NIYKIFLLQAYRFHACVIOLPFDQVRKNTLFFLGIISQASCCYAILKVNPGMTLKAS 1052  
 1000 NIYKILLQAYRFHACVIOLPFHQWKNQPTFFLRVISTASLCSYILKAKNAGMSLGA 1059  
 1053 GS---PPERAAHLCYCAFLKLAHSHVYKCLGLPLTAQKLCRKLPEATMILKAAA 1109  
 1060 GAAGPLFSEAVQWCHQAFLLKTLRHRVTVYVPLGLSLRTAQQLSRKPLPGTTLTALAAA 1119  
 1110 DPALSTDFQITLD 1122  
 1120 NPALPSDFKILD 1132

SULT 10  
 -10-054-611-225  
 Sequence 225, Application US/10054611  
 Publication No. US20030059787A1  
 GENERAL INFORMATION:  
 APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.

Harley, Calvin  
 Andrews, William H.  
 TITLE OF INVENTION: No. US20030059787A1el Telomerase  
 NUMBER OF SEQUENCES: 225  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/054,611  
 FILING DATE: 18-Jan-2002  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/854,050  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002930US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 225:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
 US-10-054-611-225  
 Query Match 59.4%; Score 3505; DB 15; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 QY 1 MTRAPRCFPAVRSLLRSRYREVWPLATFVRRLGPEGRLVQGDPKIVRTLVACQVCMHW 60  
 DB 1 MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPGQWRLVQGDPAAPRALVACQVCPW 60  
 QY 61 GSQPPADLSFQVSSLKELVARVVRQLCERNERNVLAFGPELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPPAAPSPFRQVSCLELVARVQLRCERGAKNVLAFGFALLDARGGPPFAFTSVR 120  
 QY 121 SYLPTNTVETLRVSGAMLLLSRVGDDLLVYLAHCAVLYLLVPSPCAVQVCGSPLVQICA 180  
 DB 121 SYLPTNTVETLRVSGAMLLLSRVGDDLLVYLAHCAVLYLLVPSPCAVQVCGSPLVQICA 180  
 QY 181 TTDIWPVSASVYRTPRVGRNFTNLRFLQIKSSSRQEAAPKPLALPSRGTKRHLSTSTS 240  
 DB 181 ATQAEPPPPHAGS-GPERRLG-----CERAWNHSVREAGVPLGLPAGARRRGGSASRS 231  
 QY 241 VPSAKKARCYVPRVEEGP-----HRQVLTPSGKSW-VPSPARSPVEPTAEKILSSX 292  
 DB 232 LPLPKPRGAAPEBPTPVGGSWAHQRTGRGSDRGFCVVSPAR-----PAEATSLT 286  
 QY 293 GKVDLSLS-GSYCKKPSSTLSPPRQNAQLRP-FIETHFLYSGDQGRNLNPSF 350  
 DB 287 GALSCTRHSHPSVGRHAGAPSTSRPPRPWDTPCPVVAETKHFYSSGD-KEQLRPSF 345

351 LLSNLOPLTGARLVELIFLGSPPRTSGPLCRTHRLSRRYWQVRPLFQOLLVNHAEQY 410  
346 LLSLRPSLTGARLVEIFLGSPPRPMPTGTRPLRPLPQRYWQVRPLFLELGNHAQCY 405  
411 VLLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRLVQLLRQSSPWQY 462  
453 GFLRACLCVVSASLWGTNRHRRFFKNLKFISLKGKYLQELMMWKMVEDDCHWLRS 512  
463 GFVRACLRLVPPGLWGSRRHRRFFLNTKFKFISLGHAKLSQELTWKVSVRDCAWLR 522  
513 SPGKDRVPAAEHRLRERLTLATFLWMDTVVQLLRFFYITESTFQKRLFFYKSVMS 572  
523 SPGVCVPAAEHRLREELTAKFLHMLSVVVELLRFFYVTTTQKRLFFYKSVMS 582  
573 KLSQIGVQHRLRVLRELSEVVRHODTWLAMPICRLRPIPKPGLRPIVNMYSMT 632  
583 KLSQIGIRQLKRVQLRELSEAEVRQREARFALLTSRLRPIPKPGLRPIVNMYSMT 642  
633 RALGRKQAOHFTORLKTFLSMANVETKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
643 RTFRREKRAELTSRVKALFSLVRYERARRPGLUGASVIGLDDIHRANRTFVLVRADP 702  
693 TPRMYFVKADVTGAYDAIPQGLVEVVMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
703 PPELYFVKVDTGAYDTIPQDLRTEVIAIIPK-PONTYCVRVAVVQKAAHGHVKAFFS 761  
753 QVTTLSLQPYMGQFLEKLODSASALRNSVIEQSIKNNESSSLDFELHRLHSVVK 812  
762 HVSTLTLDQPYMRQFVAHQET--SPURDAVIEQSSSLNEASSGLDFELRWMCHAVR 819  
813 IGDRCYTQCGIPQGSSTLLCSLCFGDMENKLFABEVORDGLLRFDVLLVTPHLDQ 872  
820 IRGKSVYQCGIPQGSSTLLCSLCYGMENKLFAGIRRDGLLRVDDFLVTPHLDQ 879  
873 AKTFELTLVHGVPEYGCMLNQLKTVNPEVPTLGGAAVYQLPALCLFWCCGLLDTOT 932  
880 AKTFELTLVHGVPEYGCMLNQLKTVNPEVPTLGGAAVYQLPALCLFWCCGLLDTOT 939  
933 LEVFCDSYAGTISKTSLTQSVFKAGTWRNKLKSLVLRKCHGLFLDLQVNSLQTVCI 992  
940 LEVQDSYSSYARTSIRASLTNRGFKAGNNRKLFGVLRKCHSLFLDLQVNSLQTVCI 999  
993 NIYKFLQAVRHACVTLQFPDQVRNKLTFGLIGTSSQASCCYAILKVNPGMTLKAS 1052  
1000 NIYKILLQAVRHACVTLQFPDQVRNKLTFGLIGTSSQASCCYAILKVNPGMTLKAS 1059  
1053 GS---FPPAAHLCVQAFLLKLAHSAHVYKCLLGLRLTAQKLCRLPEATWTILKAAA 1109  
1060 GAAGLPLPSEAVQWUHQAFLLKLTNRVYVYVPLGLSLRTAQTQLSRKLPGLTTLTALAAA 1119  
1110 DPALSTDFQTLID 1122  
1120 NPALPSPDKTILD 1132

## RESULT 11

S-10-105-963-2  
Sequence 2, Application US/10105963  
Publication No. US2003006818A1  
GENERAL INFORMATION:  
APPLICANT: Geron Corporation  
APPLICANT: Denning, Chris  
APPLICANT: Clark, A. John  
TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
TITLE OF INVENTION: Recombination  
FILE REFERENCE: 731/002  
CURRENT APPLICATION NUMBER: US/10/105,963  
CURRENT FILING DATE: 2002-03-21  
PRIOR APPLICATION NUMBER: US 60/277,811

PRIOR FILING DATE: 2001-03-21  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 1132  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-105-963-2

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPRCFAVRSLLRSRYREVWPLATFVRLRPGEGRRVLPQDPKIVYRTLVAAQCLVCMHW 60  
DB 1 MPRAPRCFAVRSLLRSRYREVWPLATFVRLRPGEGRRVLPQDPKIVYRTLVAAQCLVCMHW 60  
QY 61 GSOPPPADLSFHQVSSKLSELVARVQVORLERNERNVLPAGFELLNEARGGPPMATSSVR 120  
DB 61 DARPPPAAPSFQVQSLKELVARVQVORLERNERNVLPAGFELLNEARGGPPMATSSVR 120  
QY 121 SYLPTVETLRYSGAMLLLSRGDDILAVLLAHCALVLLVPPSCAYQVCGSPYQICA 180  
DB 121 SYLPTVETLRYSGAMLLLSRGDDILAVLLAHCALVLLVPPSCAYQVCGSPYQICA 180  
QY 181 TTDTPSVSASVYRTPVGRNFTNLFLOQIKSSRQEAPEKLPALPSRGTKEHLSLTSTS 240  
DB 181 ATQARPPPHAS-GPRRRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
QY 241 VPSAKKACVYPRVREGEPR-----HRQVLTPTSGKSW-VSPARSPEVPTAERDLASK 292  
DB 241 VPSAKKACVYPRVREGEPR-----HRQVLTPTSGKSW-VSPARSPEVPTAERDLASK 292  
QY 232 LPUPKPRRGAAPPEPRTVPGQSWAHPTGRTGSDRGFCVVSPPAR-----PAEATSLE 286  
DB 232 LPUPKPRRGAAPPEPRTVPGQSWAHPTGRTGSDRGFCVVSPPAR-----PAEATSLE 286  
QY 293 GKVSDDLSSLS-GSVCCCKHPSSTLSLSPRQNAFQRP-PIETRHFLYSRGDQGERLNPSF 350  
DB 287 GALSGRHSHPSNGRQHAGPPSTSRPPRWDTPCPVYVETKHEFLYSSGD-KEQLRPSF 345  
QY 351 LLSNLOPLTGARLVELIFLGSPPRTSGPLCRTHRLSRRYWQVRPLFQOLLVNHAEQY 410  
DB 346 LLSLRPSLTGARLVEIFLGSPPRPMPTGTRPLRPLPQRYWQVRPLFLELGNHAQCY 405  
QY 411 VLLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
DB 406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRLVQLLRQSSPWQY 462  
QY 453 GFLRACLCVVSASLWGTNRHRRFFKNLKFISLKGKYLQELMMWKMVEDDCHWLRS 512  
DB 463 GFVRACLRLVPPGLWGSRRHRRFFLNTKFKFISLGHAKLSQELTWKVSVRDCAWLR 522  
QY 513 SPGKDRVPAAEHRLRERLTLATFLWMDTVVQLLRFFYITESTFQKRLFFYKSVMS 572  
DB 523 SPGVCVPAAEHRLREELTAKFLHMLSVVVELLRFFYVTTTQKRLFFYKSVMS 582  
QY 573 KLSQIGVQHRLRVLRELSEVVRHODTWLAMPICRLRPIPKPGLRPIVNMYSMT 632  
DB 583 KLSQIGIRQLKRVQLRELSEAEVRQREARFALLTSRLRPIPKPGLRPIVNMYSMT 642  
QY 633 RALGRKQAOHFTORLKTFLSMANVETKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
DB 643 RTFRREKRAELTSRVKALFSLVRYERARRPGLUGASVIGLDDIHRANRTFVLVRADP 702  
QY 693 TPRMYFVKADVTGAYDAIPQGLVEVVMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
DB 703 PPELYFVKVDTGAYDTIPQDLRTEVIAIIPK-PONTYCVRVAVVQKAAHGHVKAFFS 761  
QY 753 QVTTLSLQPYMGQFLEKLODSASALRNSVIEQSIKNNESSSLDFELHRLHSVVK 812  
DB 762 HVSTLTLDQPYMRQFVAHQET--SPURDAVIEQSSSLNEASSGLDFELRWMCHAVR 819  
QY 813 IGDRCYTQCGIPQGSSTLLCSLCFGDMENKLFABEVORDGLLRFDVLLVTPHLDQ 872  
DB 820 IRGKSVYQCGIPQGSSTLLCSLCYGMENKLFAGIRRDGLLRVDDFLVTPHLDQ 879

873 AKTFLSTLVHGVPEYCGMINLQKTVNFFVEPTGLGGAAPYQIPAHCLFPWCGLLDITQT 932  
 880 AKTFLRTLVHGVPEYCGVNLKTVNFFVEDEALGGTAFVQMPAHGLFPWCGLLDITRT 939  
 933 LEVFCDSYGAQTSIKTSITFQSVFRAGKTMTNKLISVLRKACHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASLTFFNKGFRAGENMRKLFGLVLRKACHSLFLDLQVNSLQTVCT 999  
 993 NIYKIFLQAYRHACVIQPPQDVRKNTLTFPLGLISSQASCCVAILKVKNPQMTLKAS 1052  
 1000 NIYKILLQAYRHACVLOLPPHQQWKNPTFFLRVISTASLCLYILKAKNAGMSLGAK 1059  
 1053 GS---FPPEAAHLCYQAFLLKLAHRSVLYKCLLGLPRTAQKLCCKLPEATWTLKAA 1109  
 1060 GAAGLPSEAVQWLCHQAFLLKLTNRVTVVPLGSLRTAQQLSKLPGTTLTALAAA 1119  
 1110 DPALSTDFQITLD 1122  
 1120 NPALPSDFKITLD 1132

SULT 12  
 -10-044-692-2  
 Sequence 2, Application US/10044692  
 Publication No. US20030096344A1  
 GENERAL INFORMATION:  
 APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Harley, Calvin  
 Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/044,692  
 FILING DATE: 11-Jan-2002  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/912,951  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-10-044-692-2

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCPAVRSLRSRVREVMPLATFVRRLGPEGRRLVQGDPKIVRTILVAOCLVCMHW 60  
 DB 1 MTRAPRCPAVRSLRSRVREVMPLATFVRRLGPEGRRLVQGDPAAPRALVAOCLVCPW 60  
 QY 61 GSCPPPADLSFHQVSSSLKELVARVVQRLCERNENRVLAFGPELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPPPAAPSPRQVSKELVARVLQRLCERGAKNVLAFGFALLDARGGPPPEAFTTSSVR 120  
 QY 121 SYLPNTVLTETLVSGAWMLLSRVGDDLLVLLAHACALYLLVPSCAYOVCGSPLYOICA 180  
 DB 121 SYLPNTVTDALRGSGAWGLLRVGDVLLVHLARCALFVLVAFSCAYQVCGPPLYOLGA 180  
 QY 181 TTDIMPSVSASYRTPRPVGRNFTNLRLFLQIKSSSQEAPKPLALPSRGTKRHLSTSTS 240  
 DB 181 ATQARPPPEAS-GPRRLG-----CERAWNHSVREAGVFLGLPAPGARRRGGSSARS 231  
 QY 241 VPSAKKARCYVPVRVEEGP-----HQVLTPTSGKSW-VPSAPSPSEVPTAEKDLSSK 292  
 DB 232 LFLPKRPRRGAPEPERTVQGGNAHPGRTGRGSCVGVSPAR-----PABEATSL 286  
 QY 293 GKVSDDLISL-GSVCCCHKPSSSTLSLSPRQNAFQLRP-FIETRHFLYSRGGDQERLNPSF 350  
 DB 287 GALSCTRSHSPSVGQHAGPPSTSRPRPMDTFCPPVYAEKTHFLYSSGD-KEQLRPSF 345  
 QY 351 LLSNLPNLTGARLVEIIFLGSRRSGCLCRTHRLSRRYQWRPLFQOLLNVHAEQY 410  
 DB 346 LLSSLRPSLTGARLVEITFLGSRPMWPGTTPRRLPRLPQRYWQWRPLFLELLGHAQCPY 405  
 QY 411 VLLRSHCRFRATANQVTDAL-----NTSPHLLMDLLRLHSSPQVY 452  
 DB 406 GVLLKTHCPLRRA---VTPAAGVCAREKPGQSVAAPEEDTDPRVLQLLRQSSHPQVY 462  
 QY 453 GFLRACLCKVWSASLWGTNRHNERFFKNLKFISLGYKGLSLQELMMKMKVEDCHWLRS 512  
 DB 463 GFVRACLRLVPPGLWGRSHNRERFLNKKFISLGHAKLSLQELTWKMSVRDCAWLR 522  
 QY 513 SPKDRVPAAEHLREERILATFLWLMQVYVQLRSFVITESTFQKRLFFYRKSVM 572  
 DB 523 SPFGVCPAAEHLREERILAKFLHLMVSVVVELLRSFFVYTETTFQKRLFFYRKSVM 582  
 QY 573 KLSIGVRQHLRVRRLRELSQEVRRHHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMTG 632  
 DB 583 KLSIGIRQHLKRVQLRELSAEVRRHREARPAALLTSRLRFIPKPDGLRPIVNDYVVG 642  
 QY 633 RALGRKQAOHFTQRLKTLFSLNLYERTKEPHLMSGSVLGNDIYRTWRAPVLRVALDQ 692  
 DB 643 RTFRREKAERLTSRVKALFSLNYERARREGLLGASVLGLDDIHRARWTFVLRVAQDP 702  
 QY 693 TPRMYFVRADVTGAYDAIPQGLVEVVAVMIRHSESTYCIQYAVVRDSDQGVHKSFR 752  
 DB 703 PPELYFVKVDVTGAYDTIPQDLTEVIASIIK-PQNTYCVRRYAVVQKAAHGVKAFKS 761  
 QY 753 QVTTLSDIQPYNGQFLKHLQSDASALRNSVVISQSI SMNESSSLDFPFLHFLURHSVK 812  
 DB 762 HVTSLTDLQPYNMQFVAHLOET--SPLRDVAVIEQSSSLNEASSGLDFVLRPMCHAVR 819  
 QY 813 IGDRCVTCQGIPOGSSSLTLLCSLCFCGDMENKFAEVQRDGLLRRVDDFLVTPHLQ 872  
 DB 820 IRGKSYVQCGIPIQSSILSTLLCSLCGDMENKLPAGIRDDGLLRLVDDFLVTPHLTH 879  
 QY 873 AKTFLSTLVHGVPEYCGMINLQKTVNFFVEPTGLGGAAPYQIPAHCLFPWCGLLDITQT 932

```

b 880 AKTFLRTLVRGVPEYGCNVNLRKTVNPFVDEALGGTAFVQMPAHGLFWCGLLDTRT 939
Y 933 LEVFCDSYGAQTSIKTSLTQSUFKAGKATWKNKLLSVLRKCHGLFLDLQVNSLOTVCII 992
b 940 LEVQSDYSYARTSIRASLTNRGFKAGRNWRKLFGLRLKCHSLFELDLQVNSLOTVCII 999
Y 993 NYIKFLQAVRHACVITQLPDDQVRNKLTFGLIGTSSQASCCYAILKYNPQMTLKAS 1052
b 1000 NYIKILLQAVRHACVITQLPDDQVRNKLTFGLIGTSSQASCCYAILKYNPQMTLKAS 1059
Y 1053 GS---FPPEAAHLCYQAFLLKLAHSHVYIKCLLGLPRLTAQKLLCRKLPEATWTILKAAA 1109
b 1060 GAAGLPBPAVOWLCHQAFLLKLRHRTVYVPLLSGLSRLTAQQLSRKLPGLTITLALAAA 1119
Y 1110 DPALSTDFQTLID 1122
b 1120 NPALPSPDKTILD 1132

```

## RESULT 13

S-10-044-539-2

Sequence 2, Application US/10044539

Publication No. US20030100093A1

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044,539

FILING DATE: 11-Jan-2002

## CLASSIFICATION DATA:

CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

## INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-044-539-2

Query Match          59.4%; Score 3505; DB 15; Length 1132;
Best Local Similarity 62.4%; Pred. No. 1.6e-316;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCAPVRSILRLGRYREWPPLATFVRRLQEGSRLLVQPGDPKIYRLVQAQCLVCMHW 60
DB 1 MPRAPRCAPVRSILRLGRYREWPPLATFVRRLQEGSRLLVQPGDPKIYRLVQAQCLVCMHW 60
QY 61 GSQPPADLSPHQVSSIKELVARVORLCERNERNVLAFGFELLNEARGGPPMATSSVR 120
DB 61 DARPPAAPSPRQVSCIKELVARVORLCERNERNVLAFGFELLNEARGGPPMATSSVR 120
QY 121 SYLPTVETLRVSGAWMLLSRVGDDLLVYLLAHALYLLVPPSCAYQVCGSPLYQICA 180
DB 121 SYLPTVETLRVSGAWMLLSRVGDDLLVYLLAHALYLLVPPSCAYQVCGSPLYQICA 180
QY 181 TTDIWPVSASVYRPTRPVGRNFTNLRFLOQIKSSSRQEAQKPLALPSRGTKXHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231
QY 241 VPSAKKARCYVPRVREGP-----HRQVLPSPGKSW-VSPARSPEVPTAEKDLSSK 292
DB 232 LPLPKRPRGAPEPERTPVQGSWAHPGRTGRGFCVVSVPAR-----PAEATSLLE 286
QY 293 GKVSDDLSS-GSVCCCKHKPSSTLSLSPPRQNAQLRP-FIETRHFLYSGDQOERLNPSF 350
DB 287 GALSGRHSHPVGRQHAGPPTSRPPHPMDTPCPVVAETKFLYSSGD-KEQLRSPF 345
QY 351 LLSNLQNLTGARLVEIIFLGRPRTSGLRTHLSRRYQWQMRPLFOQLLVNHAQCY 410
DB 346 LLSSRLPSLTGARLVEIIFLGRPRTSGLRTHLSRRYQWQMRPLFOQLLVNHAQCY 405
QY 411 VLLRSHCRFTANQCVTDAL-----NTSPPHMLDLRLHSSFPWQY 452
DB 406 GVLLKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSFPWQY 462
QY 453 GFLRACLCKVVSASLWGTNRNERRFFKNLKKTSIGKYKLSLOELMHPKVEDCHWLRS 512
DB 463 GFVRACLRRLVPPGLWGRNRRNRRFLNRTKKFTSLGKAKLSLQELTWKMSVRDCAWLR 522
QY 513 SPGKDRVPAAEHRLREILATFLFMDTVYVQLLESFFYITESTFOKNRFFYKSWYS 572
DB 523 SPGVGCVPAAEHRLREILAKFLHLMVSIVVELGRSFFYVTTTFQKNRFFYKSWYS 582
QY 573 KLSQIGVRQHLRVRRLRELSQEEVRHHQDTLWAMPTICRLRFIPKFNGLRPIVNMYSMT 632
DB 583 KLSQIGVRQHLRVRRLRELSQEEVRHHQDTLWAMPTICRLRFIPKFNGLRPIVNMYSMT 642
QY 633 RALGRRKQAOHFTQRLKTLFSLMNYERTKPHLMSSVLGMNDIYRTWEAFVLRYALDQ 692
DB 643 RTFRREKGAERLTSRVKALFSLVLYERARRRPGLLGASVLGDDIHRAMWTFVLRYAQDP 702
QY 693 TPRMYFVKADVTGAYDAIPQGLVEVAVNMIRHSESTYCIROYAVVVRDSQGVHKSFR 752
DB 703 PPELYFVKADVTGAYDTIPQDLTEVIAIHK-PQNTYCVRRYAVVQKAAHGHVKAFFS 761
QY 753 QVTLSDLOPYMGQFLKHLQDSDASALRNSVITEQISIMNNESSSLFDFFLHLSHVVK 812
DB 762 HVSTLTDLPYMRQFVAHQET--SPLRDVAVTEQSSLSNEASSGLFDVFLRFMCHAVR 819
QY 813 IGRDCTVQCQGIPOGSSLSLTLCSLCFGDMENKLPABVORDGLLLRFVDDFLVTPHLDQ 872
DB 820 IRGKSYVQCQGIPOGSSLSLTLCSLCFGDMENKLPABVORDGLLLRFVDDFLVTPHLDQ 879
QY 873 AKTELSTLVHGVPYEGCMINLQKTVVNFVPEPOTLGGAAPYQLPAHCLFFPWGGLLDTOT 932

```

880 AKTFLRLVGVPEYCVNLKRTVVNFFVEDEALGGTAFVOMPAHGLFPWCGLLDTRT 939  
 933 LEVFCYSGVQTSIKTSITFOSVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 940 LEVOSDYSSVARTSIRASLTFRNGFKAGRMERKLFVGLRLKCHSLFLDLQVNSLQTVCT 999  
 993 NIYKIFLQAYREHACVIOLEPDORVRKNLTFELGIISSQASCCVAILKVNKPGWTLKAS 1052  
 1000 NIYKILLQAYREHACVIOLEPDORVRKNLTFELGIISSQASCCVAILKVNKPGWTLKAS 1059  
 1053 GS---PPEAAHWLCYQAFLLKLAHSVYKCLLGLPLRTAQKLLCRKLPKPEATMILKAAA 1109  
 1060 GAAGLPSEAVOMLCHQAFLLKTRHRTVYVPLGSLRTAQQLSKLPGLTTLTALFAAA 1119  
 1110 DPALSTDFQTILD 1122  
 1120 NPALPSDFKTILD 1132

SULT 14  
 -10-044-692-323  
 Sequence 323, Application US/10044692  
 Publication No. US20030096344A1  
 GENERAL INFORMATION:  
 APPLICANT: Cecch, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Harley, Calvin  
 Andrews, William H.  
 TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
 THERAPEUTIC METHODS  
 NUMBER OF SEQUENCES: 335  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/044,692  
 FILING DATE: 11-Jan-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/912,951  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002600US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 323:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1154 amino acids

TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 323:  
 US-10-044-692-323  
 Query Match 59.4%; Score 3505; DB 15; Length 1154;  
 Best Local Similarity 62.4%; Fred. No. 1.7e-316; Mismatches 260; Indels 52; Gaps 13;  
 Matches 719; Conservative 122;  
 QY 1 MTRAPRCAPAVRSILKRSRYREVMPLATFVRRLLGPEGRRLVQDPDKIYRTLVVAQCLVCMHW 60  
 DB 1 MPRAPRCAPAVRSILKRSRYREVMPLATFVRRLLGPEGRRLVQDPDKIYRTLVVAQCLVCMHW 60  
 QY 61 GSQPPADLSFHQVSSLSKELVARVVRQRLCERNENVLAQFELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPPPAAPSPRQVSCLEKELVARVVRQRLCERNENVLAQFELLNEARGGPPMAFTSSVR 120  
 QY 121 SYLPNTVTETLRVSGAWMLLSRVGDDLLVLLAHACALYLLVPSCAYQVCGSPLYOICA 180  
 DB 121 SYLPNTVTETLRVSGAWMLLSRVGDDLLVLLAHACALYLLVPSCAYQVCGSPLYOICA 180  
 QY 181 TTDIWPVSASYPTRPVGRNFTNLRLFLQIKSSSRQAPKPLALPSRGTGRHLSLTS 240  
 DB 181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSSASRS 231  
 QY 241 VPSAKKARCYVPVRVEGPG-----HRQVLPSTGKSW-VPSAPSPPEVPTAEKDLSSK 292  
 DB 232 LPLPKRPRGGAPEPERTVGGGSAHGPGRTRGSDRGFCVVSPAR-----PAEATSL 286  
 QY 293 GKVSLLSLSS-GSVCKKHPSSTLSLPPRQNAFQRLP-FIETRHFLYSRGGDQERLNPSE 350  
 DB 287 GALSCTRHSHPVSGRQHHAGPPSTSRPPMDTPCPPVYARTKHFLYSSGD-KEQLRPSF 345  
 QY 351 LLSNLQPNLTGARRLVEIIFLGSRRPSGCLRTHLSRRYQWNRPLFQOLLVNHACQY 410  
 DB 346 LLSLSRPSLTGARRLVETIFLGSRRPMEGTPTRLPLFQRYWQMRPLFLELLGNHACQY 405  
 QY 411 VLLRSHCRFRRTANQQVTDAL-----NTSPHLLMDLLRLHSSPQVY 452  
 DB 406 GVLLKTHCEPLRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLLVQLLRQHSPPQVY 462  
 QY 453 GFLRACLCKVWSASLWGRTHNRERFFKNLKEFIISLGYKGLSLQELMWKMKVEDCHLWS 512  
 DB 463 GFVRACLRRLVPPGLWGRSHNRERFLNRTKKFISLGHAKLSLQELTWKMSVRDCAWLR 522  
 QY 513 SPKDRVPAAEHRLERILATFLWMDTYVQLLRGFEYITESTFQKNRLFVYRKSVMS 572  
 DB 523 SPGVGCVPAAEHRLERILATFLWMDTYVQLLRGFEYITESTFQKNRLFVYRKSVMS 582  
 QY 573 KLQSIGVROHLERVRRLRELSQBEVRHHODTWLAMPICRLRFIPKNGLRPIVNNMSYSGT 632  
 DB 583 KLQSIGIRQHLKRVQLRELSEAEVQREARPAALLTSELRFIPKPDGLRPIVNNMDYVGA 642  
 QY 633 RALGRKQAHQHTQRLKTLFSLNRYERTKPHLMGSSVLGMNDIYRTWRAFLVLRALDQ 692  
 DB 643 RTFRREKKAERLTSRVKALFSLNRYERARRPCLLGASVLGLDDTHRAWRTFVLVRADDP 702  
 QY 693 TRPMYFVKADVTGAYDAIPOGKLVVAVNMRHSESTYCIQYAVAVRDSQGVKSKFR 752  
 DB 703 PPELVFKVDVTGAYDTIPQDELTEVIAIILK-PQNTYCVRYAVVQKAAGHVYKAFKS 761  
 QY 753 QVTLTSLDQPYNGQFLKHLQSDASALRNSVVIQSSISMNESSSLDFFLFHLFURHSVK 812  
 DB 762 HVSITLTLQPYNMQFVAHLQET--SPLADAVVIEQSSSLNEASSGLEDFVLRFMCHAVR 819  
 QY 813 IGDRCVTCQGGIPQGSLSLTLCSLCFGDMENKLFQVORDGLLRLVDDFLVTPHLQ 872  
 DB 820 IKGKSVQCCGIPQGSLSLTLCSLCYCGDMENKLFQVIRDRGLLRLVDDFLVTPHLQ 879  
 QY 873 AKTFIATLHVGVPEYGCMLNLOKTVNFPFVPGTLLGGAAPYQLPAHCLFPWCGLLDTQT 932

880 AKTFLRLVRGVPEYGVNLRKTVNFPVEDEALGGTAFVQMPAHGLFPNGLLDTRT 939  
 933 LEVFCYSGVAQTSIKTSITFSQVFKAGKTMNKLLSVLRKCHGLFELDQVNSLQTVCI 992  
 940 LEVQSDYSSVARTSIRASLTFRNGFRAGRMNRKLFVGLRLKCHSLFLDQVNSLQTVCT 999  
 993 NIYKFILOAYRHACVIOPLPFDORVKNTLTFELGIISSQASCCVAILKVNPGMTLKAS 1052  
 1000 NIYKILLQAYRHACVLOLPPFHQQVWKNPTFFLRVISTATSLCYSIIKAKNAGNSLGAK 1059  
 1053 GS---FPPEAAHLCYQAFLLKLAHSHVYKLLGLPLRTAQKLLCRKLPKPEATMTILKAA 1109  
 1060 GAAGPLSEAVQMLCHQAFLLKLTREHVTVPLGLSLRTAQQLSEKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQILD 1122  
 1120 NPALPSDFKILD 1132

# 38ULT 15

3-10-044-539-323  
 Sequence 323, Application US/10044539  
 Publication No. US20030100093A1  
 GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Harley, Calvin  
 Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/044,539  
 FILING DATE: 11-Jan-2002  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/912,951  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002600US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 323:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1154 amino acids

TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 323:  
 US-10-044-539-323  
 Query Match 59.4%; Score 3505; DB 15; Length 1154;  
 Best Local Similarity 62.4%; Pred. No. 1.7e-316; Mismatches 260; Indels 52; Gaps 13;  
 Matches 719; Conservative 122;  
 QY 1 MTRAPRCPAVRSLRSRYREVMPLATFVRLGPEGRILVQDPDKIYRTLVLAQCLVCMHW 60  
 DB 1 MTRAPRCRAVRSLRSRYREVMPLATFVRLGPEGRILVQDPDKIYRTLVLAQCLVCPV 60  
 QY 61 GSQPPPADLSFHOVSSLSKELVARVQRLCERNERNVLAFGPELLNEARGGPPMAFTSVR 120  
 DB 61 DARPPPAAPSFQVSCLELVARVQRLCERGAKNVLAFGFALLDARGGPEPFAFTSVR 120  
 QY 121 SYLPNTVIELRVSGAWMLLSRVGDDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180  
 DB 121 SYLPNTVTDALRGSGANGLLRRVGGDVLVHLLARCALFVLVAFSCAYQVCGPPLYQLGA 180  
 QY 181 TTDIWPVSASVYRTPRPGVGRNFTNLRFLQOKSSSRQBPAPKPLALPSGTYKXHLSTST 240  
 DB 181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSASRS 231  
 QY 241 VPSAKKARCYPPRYVEEGP-----HEQVLPSTGKSW-VPSPARSPVEVPTAEKDLSSK 292  
 DB 232 LPLPKPRRGAAPFERIPVGGQSWAHFGRTRGSDRGFCVVSAR-----PAEATSL 286  
 QY 293 GKVSIDLSS-GSVCKKHPSTSLSPRONAFQLRP-FIETRIPLYSRGGQRLNPSF 350  
 DB 287 GALSGRHSHPSVGRQHGAGPSTSRPPMDTPCPVYAEKHEFLYSSGD-KEQLRPSF 345  
 QY 351 LLSNLQNLTGARLLVEIIFLGSRRPTSGPLCRHLSRYWQMPLEFQOLLVNHAEQY 410  
 DB 346 LLSLSRPSLTGARRLVETIFLGSRRPMPGTPRLPRLPQRTWQMRPLELELGHNAQCPY 405  
 QY 411 VRLLRSHCRFTANQQVTDAL-----NTSPHMLDLLHSHSPWQVY 452  
 DB 406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEBEDTPRLVQLLRQHSSPWQVY 462  
 QY 453 GFIRACLCVKVVSASLWGRHNERRFKNLKXFIISLGYKGLSLOELMWMKVEDCHWLS 512  
 DB 463 GFVRACLRLVPEGLWGRHNERRFKNLKXFIISLGYKGLSLOELTWKNSVRDCAWLRR 522  
 QY 513 SPGKDRVPAAEHRLRERILATLFWLMDTVVQLLRSFFYITESTFQKNRFFYRKSVMS 572  
 DB 523 SPGVGCVPAAEHRLRERILATLFWLMDTVVQLLRSFFYITESTFQKNRFFYRKSVMS 582  
 QY 573 KLOSIGVROKLERVRLRELSQBEVRHQDTWLAMPICRLRIPKNGRLPIVNNYSYNGT 632  
 DB 583 KLOSIGTRQKLRVQRLRELSQBEVRHQDTWLAMPICRLRIPKNGRLPIVNNYSYNGT 642  
 QY 633 RALGRKQAOHFTORLKTFLFMINVERTKHPHMGSSVLGMNDIYRTWRAFVLRVRLDQ 692  
 DB 643 RTRERKRAERLTSRVKALFVSVLNVVERARRPGILASVLGLDDIHRANRTFVLRVRAQDP 702  
 QY 693 TPRMYFYKADVTGAYDAIPQGLVEVAMNHSSTYCIQYAVVRDSDGQVHKSPRR 752  
 DB 703 PPFLYFYKADVTGAYDTIPQDRLTEVIAIIR-PONTYCVRYAVVQAAHGHVRKAFKS 761  
 QY 753 QVTTLSLQPYMGQFLKHLQDSDASALRNSVVEIOSISMNNESSSLFDFFLHFLRHSVVK 812  
 DB 762 HVSTLTDLPYMEQFVAHQEI--SPLRDAVVIEGSSLSNEASSGLFDVFLFMCHAVR 819  
 QY 813 IGRCYTCCQGI PQGSSLSLTLCSLCFGDMENKLFABVORDGLLRFVDDFLLVTPHLDQ 872  
 DB 820 IGRKSYVCCQGI PQGSSLSLTLCSLCFGDMENKLFABVORDGLLRFVDDFLLVTPHLDQ 879  
 QY 873 AKTFLSTLVHGVPEYGVNLRKTVNFPVEDEALGGTAFVQMPAHGLFPNGLLDTRT 932

```

880 AKTFRLTVRGVPEYCCVNLKRTVVNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 939
933 LEVFCDYSGYAQTSIKTSLTFSVPKAGTKMENKLLSVLRKXCHGLFLDLQVNSLQTVCI 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
940 LEVQSDYSYARTSIRASLTFRNGFKAGRNMRKLFGLVLRKXCHSLFLDLQVNSLQTVCT 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
993 NIYKIFLQAYREHACVIOLPEDORVRKULTFEFLGISSQASCCYAILKVXNPGMTLKAS 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1000 NIYKILLQAYRFHACVLOLPFHQQVQWKNPTFFLRVISDTASLCYSIILKAKNAGMSLGAK 1059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1053 GS---FPPEAAHWLCYQAFLLKLAHSHVIYKCLLGLPLRTAQKLLCRKLPEATMTILKAAA 1109
    | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 GAAGPLPSEAVQWLCHQAFLLKLTTRHRTVYVLLGSLRLTAQTQLSRKLPGTTLTALAAAA 1119
    | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1110 DPALSTDFQTILD 1122
    : ||| : ||| |||
1120 NPALPSDFXTILD 1132

```

arch completed: February 4, 2004, 14:17:56  
b time : 50 secs



GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.  
 protein - protein search, using sw model  
 n on: February 4, 2004, 14:09:14 ; Search time 27 Seconds  
 (without alignments)  
 3996.343 Million cell updates/sec  
 tie: US-09-042-460-2  
 effect score: 5901  
 quence: 1 MTRAPCPAVRSLLRSRYE.....TLKAAADPALSTDFQILD 1122  
 oring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

arched: 283308 seqs, 96168682 residues  
 tal number of hits satisfying chosen parameters: 283308

num DB seq length: 0  
 ximum DB seq length: 2000000000  
 st-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

atabase :  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

ult	No.	Score	Query Match	Length	DB ID	Description
1	3505	59.4	1132	2	T03844	telomerase catalyt
2	777.5	13.2	1123	2	T51517	telomerase reverse
3	624	10.6	989	2	T03838	telomerase catalyt
4	426.5	7.2	1132	2	T31107	telomerase reverse
5	401.5	6.8	1117	2	T14891	telomerase (EC 2.7
6	341	5.8	884	2	S53396	telomerase catalyt
7	133.5	2.3	561	2	T20392	hypothetical prote
8	131.5	2.2	330	2	E98119	transposase, uncha
9	131	2.2	778	2	I38487	tastin - human
10	129	2.2	502	2	T24227	hypothetical prote
11	125	2.1	1529	2	T02730	RNA-directed DNA p
12	124	2.1	4540	2	T30838	cytoplasmic dynein
13	121	2.1	2088	2	E71436	hypothetical prote
14	120.5	2.0	1646	2	T20740	hypothetical prote
15	116.5	2.0	544	2	E11216	reverse transcript
16	116	2.0	881	2	E82097	protein-P-II uidity
17	115.5	2.0	2157	2	S71461	proline-rich prote
18	114.5	1.9	1587	2	AB2012	hypothetical prote
19	111.5	1.9	1415	1	EDBEGA	immediate-early pr
20	111.5	1.9	2606	2	T03159	large tegument pro
21	110	1.9	907	2	A24938	hypothetical T2 pr
22	109.5	1.9	710	2	T41586	hypothetical serin
23	109.5	1.9	1214	2	T00356	hypothetical prote
24	109.5	1.9	2493	2	T40540	hypothetical prote
25	109.5	1.9	3147	2	T21328	hypothetical prote
26	108.5	1.8	351	2	S50754	hypothetical prote
27	108.5	1.8	1147	2	T42627	ADP-ribosylation f
28	108	1.8	880	2	S49627	regulatory protein
29	108	1.8	1474	2	B85188	retrotransposon li

30	107.5	1.8	600	2	AS2570	hypothetical prote
31	107	1.8	510	2	H84824	zn/Spm-like transp
32	107	1.8	949	2	T38543	hypothetical prote
33	107	1.8	1003	2	E85432	hypothetical prote
34	107	1.8	1121	2	A82809	exodeoxyribonuclea
35	107	1.8	1456	2	T01397	LTR gag/pol polypr
36	106.5	1.8	1256	2	C71436	probable resistanc
37	106	1.8	584	2	S40013	hypothetical prote
38	106	1.8	877	2	T43449	hypothetical prote
39	106	1.8	1017	2	T48452	hypothetical prote
40	106	1.8	1115	2	A45761	Ca2+-transporting
41	106	1.8	2528	2	T20719	hypothetical prote
42	105.5	1.8	745	2	H85048	probable transposo
43	105.5	1.8	811	2	T25166	hypothetical prote
44	105.5	1.8	3998	2	S57437	genome polyprotein
45	105	1.8	1493	2	T10757	MAP Kinase kinase

ALIGNMENTS

RESULT 1

T03844  
 telomerase catalytic chain - human  
 N;Alternate names: telomerase reverse transcriptase  
 C;Species: Homo sapiens (man)  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
 C;Accession: T03844  
 R;Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner,  
 Science 277, 955-959, 1997  
 A;Title: Telomerase catalytic subunit homologs from fission yeast and human.  
 A;Reference number: Z15111; MUID:97400623; PMID:9252327  
 A;Accession: T03844  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1132 <NAK>  
 A;Cross-references: EMBL:AF015950; NID:g23330016; PIDN:AACS1672.1; PID:g2330017  
 A;Experimental source: kidney  
 C;Genetics:  
 A;Gene: TRT  
 A;Map position: 5p

Query Match	59.4%	Score	3505	DB 2	Length	1132			
Best Local Similarity	62.4%	Pred. No.	1.9e-252						
Matches	719	Conservative	122	Mismatches	260	Indels	52	Gaps	13
QY	1	MTRAPCPAVRSLLRSRYEYVPLATFVRRIGPEGRRLVQDPKIVRTLVACLVCMHW	60						
Db	1	MTRAPCPAVRSLLRSRYEYVPLATFVRRIGPGQWRLVQRGDPAPAFRALVACLVCPW	60						
QY	61	GSQPPADLSFHQVSSSLKELVARVQRLCERNERNVLAFFELLNEARGGPPMAFTSSVR	120						
Db	61	DARPPAAPSPQVSCLEKELVARVQLQRCERGAKNVLAFFGALLDQARGPPEAFTTSVR	120						
QY	121	SVLPNTVTLRVSGAWMLLSRVGDDLLVLLAHCALVLLVPSCAYVCGSPLYQICA	180						
Db	121	SVLPNTVTLRVSGAWMLLSRVGDDLVLLHLLARCALFVLVAPSCAYVCGGPLYQLGA	180						
QY	181	TTDIPWSVASRPRFPVGRNFTNLRPLQOIKSSSRQEAAPKPLALPSRGTKRHLSTSTS	240						
Db	181	ATQARPPPHAS-GPRRELG-----CERAWNSVREAGVPLGLPAGARRRGGSSASRS	231						
QY	241	VPSAKARCPYPRVTEGP-----HQVLTPTSGKSW-VPSAPSPPEVPTAKDLSSK	292						
Db	232	LFLPKRPRGAPPEPTFVGQSWAHFGRTGRGSDRGFCVSPAR-----PABEATSL	286						
QY	293	GVKSDLSLGS-GSVCKHKPSSTLSLSPRQNAFLRP-FIETRHLYSRGDGQERLNPSF	350						
Db	287	GALSGTRHSPSGVGRQHAGPPTSRPRPMDTFCPPVYAEKHFLLYSSGD-KEQLRPSF	345						
QY	351	LLSLQPNLTGARLVEIFIGSRPRTSGPLCRTHRLSRYYQWRPRLFCQLLVNHAEQY	410						
Db	346	LLSLRSLPTGARLVETIFLGSRPWMEGTPRRLPRLPQRYWQWRPFLFELGNHACQPY	405						

Y 411 VLLRSHCHFRFRTANQOVTAL-----NTSPPHLMDLLRLHSSPQVY 452  
b 406 GVLNTHCHPLRAA---VTPAAGVCAREKPOGSAAPBEEDTPRLVOLLRQHSPPQVY 462  
Y 453 GFLRACLCKVVSASLWGTNRHNERFFKNLKFISLGVKGLSLOELMMKMKVDDCHWLRS 512  
b 463 GFLRACLCKVVSASLWGTNRHNERFFKNLKFISLGVKGLSLOELMMKMKVDDCHWLRS 522  
Y 513 SPGRDVPAAERHLRERILATFLFWMDTYVYVOLLRFFYITESTFOKNLFFYRKS VMS 572  
b 523 SPGVGCVPAERHLRERILATFLFWMDTYVYVOLLRFFYITESTFOKNLFFYRKS VMS 582  
Y 573 KLOSGVROHLERVLRLBELSQEVRHODTWLAMPICELREIPKPNGLRVLVMSVSGT 632  
b 583 KLOSGVROHLERVLRLBELSQEVRHODTWLAMPICELREIPKPNGLRVLVMSVSGT 642  
Y 633 RALGRKQOQHTQRLKTLFSLMNTYRTHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
b 643 RYFREKRAERLTSVKALFSLVNYERARPCLLGASVLGDDIIRAWRTFLVRRAQDP 702  
Y 693 TRPMYFVADVTGAYDAIPQGLVEVVMNRHSESTYCIROYAVVRDSDGQVHKSFR 752  
b 703 PPELYFVKVDVTGAYDTIPQDLTEVIAIIF- PONTYCVARYAVWQAAHGHVRAKFS 761  
Y 753 QVTTLSLOPYMGQFLKLODSASALRNSVVEIOSISMNESSSLFDFLFLRHSHVK 812  
b 762 HVSITLTDLPYMQFVAHLOET--SPLRDVVIQSSSLNEASSGLFDVFLRCHHAVR 819  
Y 813 IGRDVCYQCGIPQSSSLTLCSCFCGDMENKLPFAVORGLLRLFRVDDPLVTPHLQ 872  
b 820 IRGKSYVQCGIPQSSSLTLCSCFCGDMENKLPFAVORGLLRLFRVDDPLVTPHLQ 879  
Y 873 AKTFELTLVHGVPEYCMNLCKTVNPPVPEGTGGAPYOLPAHCLFPMCGLLDTOT 932  
b 880 AKTFELTLVHGVPEYCMNLCKTVNPPVPEGTGGAPYOLPAHCLFPMCGLLDTOT 939  
Y 933 LEVFCDSYGAOTSITSLTFSQVFKAGTMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
b 940 LEVQSDYSVARTSIRASLTFRNFKAGRMERKLPGLVLRKCHGLFLDLQVNSLQTVCT 999  
Y 993 NIYKIFLQYRPHACVQLPQDQVRKLTPLGIISSQASCCYAILKVNPGMTLAKS 1052  
b 1000 NIYKIFLQYRPHACVQLPQDQVRKLTPLGIISSQASCCYAILKVNPGMTLAKS 1059  
Y 1053 GS---PPPEAAHMLCYQALLKLAHSHVYKCLLQPLRTAQKLLCRKLPEATMTILKAA 1109  
b 1060 GAAGPLFSEAVQNLCHQAFLLKLTTRHVTYVPLGSLRTAQKLLCRKLPEATMTILKAA 1119  
Y 1110 DPALSTDFQILD 1122  
b 1120 NPALPSDFKTILD 1132

RESULT 2  
151517  
c:elomerase reverse transcriptase - Arabidopsis thaliana  
v:Alternate names: protein F5E19\_190  
s:Species: Arabidopsis thaliana (mouse-ear cress)  
c:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
a:Accession: T51517  
s:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Men  
submitted to the Protein Sequence Database, August 2000  
a:Reference number: Z25394  
a:Accession: T51517  
a:Status: preliminary  
a:Molecule type: DNA  
a:Residues: 1-1123 <SAT>  
a:Cross-references: EMBL:AL391147  
a:Experimental source: Cultivar Columbia; BAC clone F5E19  
c:Genetics:  
a:Map position: 5  
a:Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2

A&gt;Note: F5E19\_190

Query Match 13.2%; Score 777.5; DB 2; Length 1123;  
Best Local Similarity 24.0%; Pred. No. 2.7e-49;  
Matches 289; Conservative 209; Mismatches 435; Indels 269; Gaps 42;

QY 38 LVQPGDPKIVRTLVAOCLVCMHNGSQPPADLSFHQVS--SLKELVARVVQRL--CERN 92  
DB 60 LIRDDDIHVRKLLHRCFVVLH--EQTPL-LDPSFSTSWSQREIVERIIEHQSGCDC- 115  
QY 93 ERNVLAFFGELLNEARGPPMAFTSSVRSYLPNTVITLRSVGMMLLSRVGDDLLVYL 152  
DB 116 -QNVICARYD-----KYDQSSPILELTSSWEFLIKRVGHDVYVYL 156  
QY 153 LAHCALYLLVPPSCAYVCVCSPLVQICATTDIWPVSASVSRPVRPGRNFTNLRLFLQIK 212  
DB 157 LQQTISIFLLPKKKGQVSGPL-----CIKHRTLSVHEN----- 192  
QY 213 SSSRQEAAPKPLALPSPGCTKRH-LSLTSTSVPSAKARCYVPVRVEEGPHROVLPSPGKS 271  
DB 193 KRKDDNVQP-----PTKQWLSSAVDDCPKDDSATITPVGEDVDQHKRKTTKRSRI 246  
QY 272 WVPSPASPEVPTRAKDL-----SSKGVS-----DLSLSSVCCKHKPSTSL 316  
DB 247 YLKRKRQKRVNFKKVCNAPCITPTSTNGKVSFGNDEMNLHIGINGSL-----TDFV 298  
QY 317 SPQRNAFQLRPFTETRHFLYSRGDQGERLNPFLLSNLPNLGTARRLVEIIF----- 370  
DB 299 K-----QAKVKENKFKGLSETYSVIPPHILKTLRPNCSKLLMNHIFGBVVM 351  
QY 371 -----LGSPPRTSGPLCRTHRLSRRYWOMRPLFOQLLVNHAECQVYLLRSHC----- 418  
DB 352 STTPSHGKGNCP--SGSICLYHSLK--SLKNLIGTKTSHLK---MLLDKHCVPVLL 402  
QY 419 -----RERTANQOVTALNTSPH-----LMDLLRLHSSPQVYGLR 456  
DB 403 QEDALKSTTSQSRRAKADL-----PHGSSSQTGPKPCPSVEERKLYCTNDQVSVFIW 458  
QY 457 ACLCKVWSASLWGTNRHNERFFKNLKFISLGVKGLSLOELMMKMKVDD----- 506  
DB 459 AICRYIVPESLGTTHQWVLRKNIAFWVSRRNEKCTVQNFLEHVKPSPDFPFARKELC 518  
QY 507 C-----HMLRSSPGKDRVPAEHLRLERILAT-FLFWMDTYVYVOLLRSPFYITESTFOKN 561  
DB 519 CMVNGHELOE-----SIRTOQMLCTKWISLWLEIVKGLVHFNFYATESQGGRL 569  
QY 562 RLFFYRKSWSKLOSIGVROHLERVLRLBELSQEVRHODTWLAMPICRLRFPKPNGLR 621  
DB 570 NIYYRKESWERLISKELSKALDGYVLVDDAEASRRK-----LSKPRFLPKANGVR 622  
QY 622 PIVNMSYMGTRALGRKQAOHFTORLKTLPFMLNYERTKPHLMGSSVLGMNDIYRTWR 681  
DB 623 MYLDFSSSR-----SQSLRDTHAVLKQIQKEPDLVSGSVDFHDDFVFNLC 669  
QY 682 AFVLRVRALD-QTPRMVYKADVTGAYDAIPQGLVEVVMNRHSESTYCIROYAVV-- 738  
DB 670 PYLIHLRSQSGELPLFVAVADVFKAFDSVDQGLLHVIOFPLK---DEVILNRCLVCC 726  
QY 739 -RDS-----QGVHKSFRQVTTLSLQPYMGQFLKLODSASALRNSVVIQSIS 790  
DB 727 GKRSNWNKILVSSDKNSNFSRTSTV---PY-----NALQ-STVVDKGEN 768  
QY 791 NNESSSSLDFLFLHSHVVKIGDRVCYQCGIPQSSSLTLCSCFCGDMENK----- 846  
DB 769 HVRKLDLVMVIGNMLKNMLQLOKSFYVQIAGIPQGRILSSLLCCFYVGHLERTLIYPF 828  
QY 847 FAEVQD-----GLLLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGC 889  
DB 829 LEEASKVSKESGREGELIIPTSYKLLRFIDYLFVSTSRDQASSFYHLKHFQKDYNC 888  
QY 890 MINLOKTVNFPVPGTLGGAPYOLPAHCLF-----PWCGLLDDTQTLVEFCDYS 940  
DB 889 FMNETKFCINFE-----DKHEHRCSSNRMFVGDGVPFVRWVTGLLINSRTTEVQVDYT 941

941 GYAQTSTKTSITQSVFVAKGKTNRNKLVLRLKCHGLFGLDQVNSLQVTCVINYKIFLL 1000  
 942 RYLSGHSITSTFSAWQKPNVRNLRQKLCYFLVPKCHPILFDSNINSGEIVRLNIYQIFLL 1001  
 1001 QAYRFHACVIQLPDORVRKNLTF-FLGISSQASCCYAILKXVPGMTLKASGSFPP-- 1057  
 1002 AAKPHCYVEVERFWKHLHQTLFKITI-----SVYMFRLINRVRINTGSRFVVL 1056  
 1058 ----EAAHWLICYQAFLLKLAHSVYIKLGLPLRTAQKLLCRKLPEATMTILKAADPAL 1113  
 1057 KLYKEEIVMLGLDAYQVLKKNRSYRMLLIYKSA--LSKHSLSQQLSSSELRYATDRSN 1114  
 1114 ST 1115  
 1115 SS 1116

SULT 3  
 3638  
 lomerase catalytic chain - fission yeast (Schizosaccharomyces pombe)  
 Alternate names: telomerase reverse transcriptase 1  
 Species: Schizosaccharomyces pombe  
 Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Dec-2000  
 Accession: T03838; T03839; T40085  
 Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner, J.  
 ience 277, 955-959, 1997  
 Title: Telomerase catalytic subunit homologs from fission yeast and human.  
 Reference number: T15111; MUID:97400623; PMID:9252327  
 Accession: T03838  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-989 <NAK>  
 Cross-references: EMBL:AF015783; NID:92340167; PIDN:AAC49803.1; PID:G2340169  
 Experimental source: strain 972h(-)  
 Accession: T03839  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-524,526-989 <NA2>  
 Cross-references: EMBL:AF015783; NID:92340167; PIDN:AAC49802.1; PID:G2340168  
 Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
 blitted to the EMBL Data Library, March 1998  
 Reference number: Z21904  
 Accession: T40085  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-524,526-989 <LYN>  
 Cross-references: EMBL:AL022299; PIDN:CAAL18391.1; GSPDB:GNO0067; SPDB:SPBC29A3.14c  
 Experimental source: strain 972h-; cosmid C29A3  
 Genetics:  
 Query Match 10.8%; Score 624; DB 2; Length 989;  
 Best Local Similarity 23.7%; Pred. No. 6e-38;  
 Matches 255; Conservative 194; Mismatches 381; Indels 244; Gaps 46;  
 76 SLKELVARVVQRLCERN---ERNVLAFFGELNE-ARGGPPMAFTSSVRSYLPNTVIETL 131  
 83 SOSELIANVVVKQFDSFERRNLLMKGSMNHEDFRAMVNGVQNDLVSTFPNVLISIL 142  
 132 RVSGAMMLLSRVGDDLLVYLAHCAALYLLVPSCAYQVCGSPLYQICATTDIWPVSAS 191  
 143 E-SKNWOLLLEIGSDAMHYLLSKGIFEPALPNDNYLQISGIPLFK----- 187  
 192 YRTPRPGRNFTNRLFLQIKSSRSQEAPEKPLPSGKTHLSLTSTSV-----PS 243  
 188 -----NNV--FEETVSKKRTTSTSTQKSAKEVSNWSISISRFSIFYRSS 234  
 244 AKKACVYVPRVEGPHRQVLPTPQSKSWVPSPARSPEVPTAEKOLSKGKVSLSLS-G 302

235 YKKFK-----QDLYFNHL 247  
 303 SYCCKHKPSSTSLSPPRQ---NAFQLRPFIETRHFLYSGDQERLNFSLSNLQPN 358  
 248 SICDRNTVHMWLQWIFPQFGLINAFQVKL-----HKVPLVQSSTVVP- 292  
 359 LTGARLVEIIFLGRSPRTSGPLCETHLSRRYQWMBPLFQQLLVNHAECQVYVLLRSHC 418  
 293 ----KLLKY-----PIEOTAKELHRIS-----LSKYNH-YCFYI- 325  
 419 RERTANQOYTDALNTSPHLMDLRLHSSPMQVYGFACICLKCVVASLWGTNRNRRF- 477  
 326 ----DTHDE-----KILSYSLKPNQVAFSLRSILRVFPKLING---NQRIFE 367  
 478 --FKNLKFIISLGKYGKLSLOELAMKMKVEDCHW-----RSSPGKDRVPAAEHLRERIL 531  
 368 IILKOLETFLKLSRYESFSLHYLMSNIKISIEIWLVLGKRSNA---KMCISDFEKKQIF 424  
 532 ATFLFWMIDTVVQLLRSPFYITESTFQKNLRFVFKSVMSKLSQISGVQHLERVRL--- 588  
 425 AEFYWLNSFIPILOQFFYITSSDLRNTVFRKDIWLL-----CRFFITSMKEAF 480  
 589 RELSQEVRHHQDTW-LAMPICRLRFIPKPNGLRPIVMSYSMTGTRALGRKQAOHT-Q 646  
 481 EKINENNVR--MDTQKTLTPPAVIRLLPKKNTFRLLITNLKRFLIKQMGSKMLVSTNQ 538  
 647 RLKTLFESMLNYERTKPHLMG--SSVLGMN-DIYR---TWRAFVLVRALDCTPRMYFVK 700  
 539 TLRVASILK-----HLNEESSGIPFNLEVYMKULTFPKOLLKHRMFGK--KTYFVR 589  
 701 ADVTGAYDAIPQGLVEVVMNIRHSESTYCIROYAVVRDSQGVHKSFRQVTTLSDL 760  
 590 IDIKSCYDRIKQDILMFRIVKVKLXDP--FVIRKYATTHTSD--RATKNFVSEAFSYFDM 646  
 761 QPY--MGQELKHLQSDASALNSVVFQSISSMNESSSLDFDFLHFLRHSVYKIGRCY 818  
 647 VPFEKVVOLLS-MKTS-----TLFVDVDYVTKSSSEIFPKMLKHLSGHIVKIGNSQY 699  
 819 TQCOGIPQSSSLTLLCSLCFQDMENKLFABVQDQDGLLLRFVDDFLVTPHLDQAKTFL 877  
 700 LQKVGIPQGSILSSFLCHFYMEDLIDEVLSFTKKKGVLLRVVDDFLFITVKKDAKKFL 759  
 878 STLHGVGPBYGCKINLQTVNVPVEPQTLGGAAPYQLPAHCLFPWCGLLDQTLVLF- 936  
 760 NLSLRGFEKHNFTSLEKTIVNFNSNGIINN--FFNESKRMPPFPFSGFVMSRLDTLL 817  
 937 ---CDYSYGAQTSIK-TSLTFQSVFKAGTMRNKLKSVLRKCHGLFGLDQVNSLQVTC 991  
 818 ACPKIDEALFNSTVELTKHMKGSFFY--KILRSSLASFAQ-----VFIDITHNSKFNESC 870  
 992 INIYKI-----FLQAYRFHACVIQLP-----FPQVRKNLTFFLGISSQASC 1035  
 871 CNIXELGYSMCRAQAVLKRMKDFIPQRMFETDLLNVIGRKIKWKLAEILGYTSRRFLS 930  
 1036 CYAILKVKXPGNT--LKASGSFPPPEAAHMLCYQAFLLKLAHSVYIKLGLPLR 1087  
 931 SAEVKNVFLCLGRDGLKPSFKYHP-----CFEOLTYQFOSLTDLIK-----PLR 974

RESULT 4  
 T31107  
 telomerase reverse transcriptase - Oxytricha trifallax  
 C;Species: Oxytricha trifallax  
 C;date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
 C;Accession: T31107  
 R;Bryan, T.M.; Sperger, J.M.; Chapman, K.B.; Cech, T.R.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 8479-8484, 1998  
 A;Title: Telomerase reverse transcriptase genes in Tetrahymena thermophila and Oxytricha  
 A;Reference number: Z20985; MUID:98337940; PMID:9671703  
 A;Accession: T31107  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1132 <BRY>

```

;Cross-references: EMBL:AF060230; NID:G3342795; PID:G3342796; PIDN:AAC39163.1
;Genetics:
;Gene: TERT

Query Match      7.2%; Score 426.5; DB 2; Length 1132;
Best Local Similarity 19.4%; Pred. No. 3.8e-23;
Matches 194; Conservative 204; Mismatches 418; Indels 185; Gaps 37;

Y 75 SSKELVARVORLCERNERNVL-AFGELLNARGGP---PMAFTSSVRSYLPNTVI-- 128
b 153 NYIDMISKVIELLIENKLNKLOTFGKLVNNEFGNQHLGMQOQDSSHNSNEFWKC 212
Y 129 -----ETLRVGAWMLLSRVGDDLLVYLLAHACALYLLVPPSCAYOVCGSPVQICATT 182
b 213 DYINLNKQCMITKWEKYFYLDHFLPHIKEYWIFLKRDESLVIGSTNIP--CYLN 270
Y 183 DTWPSVASYRTRVGRNFTNRLFLQOIKSSQBEAPKPLALPSRTKXHLSTSTSV 242
b 271 EKLGRLOAA--YEGFNKAAN-----SAAQGSNPEANDLISAPQRK-----INTAI- 315
Y 243 SAKARCYVPVRVSGPHRQVLPTPSGKSWVPSPARSPVPTAEKDLSSGKYSD-LLSLS 301
b 316 VMKTHKYNKAAD-----SYLTN-----QEGFWDQDKRN 348
Y 302 GSVCKKHPSSTLLSPRONAFQLRPFIETRHLYSRGDQGERLNPSFLLSNLQPNLTG 361
b 349 RLIFYCAH-----QNR-----FQGHILNSKTLSSQOIR-----DNRYKEVFG 385
Y 362 ASRLVEIIFLGSRPRTSGPLCETHLSRRYQWMLPQOLLVNHAEQYVRLRSHC--- 418
b 386 FRV-----RSLKG-----KWSIIEQVIVNOKKDFKYLSKNCPLP 424
Y 419 -RFTANQOVTALNTS---PPHMLDLRLHSSPWQVYGLFRLACLCKVVSASLWGRHN 473
b 425 ENKNLKKSFLEDAVSGELRQVPRQLFECQDORQISNLTFFVANVPKFNLEK-N 483
Y 474 ERFPKALKKFTSLGKYKLSLOELMWQKVEDCHWLRSSEPKORVPAAEHRLERILAT 533
b 484 KXIFNKOMLQFVFRFESFTKISILNFRYNEVSWL-SFKCKDENKKFFMNEHVFVK 542
Y 534 FLFLMDTVVOLLRSFFVITESTQKRLFFYKSVMSKQSIGVROHLERVLRLS 593
b 543 VLKWPEDLAILMRCYFISTEKAKEYQRIFYRKNINWMLRSLIDDLKQ-NLKQVEK 601
Y 594 EVRRHQDTWLAMPICRLRIFPKPGLPPIVNMVSMGTRALGRKQAOHFTORLTKFS 653
b 602 KEMRIFCESQNPAP-GKRLIPKGTFRPIMTFRNKIPNQ-VGKFSQRMITNNKLQTAHM 659
Y 654 ML-NYERTKPHLMGSSVLGMNDIVRTWRAFLVRALDQTPRMVYFKADVGTGAYDAIPQ 712
b 660 MLKNLKSMMFKSFGFAVFNVDIMKRYENFVQWKQIN-SFKUIYFAMDIKCYDNVD- 717
Y 713 GKLEVEVANMIRHS---ESTYCIROYAVVRDSDGOVHKS-FRQVTTLSDLQPMGQFL 768
b 718 ---CERVNFLQSDLMDKPEFILNTEFLVKRNNIIVERSNFRK---LPKQVYFKFQ 770
Y 769 K-----HLQD---SDASALNSVIEQSI SMTESSSSSLFDFFLFLRHSVVKI 813
b 771 KIGIDSSYPTLFEILEDEFDNLN-MKSTIIIVEGQRKFPKNDLLOQVLKIKONNVTF 829
Y 814 GDRCTYQCGIPQSGSLSTLLCSLCFGDMENKLPFAEVORD-----GLLRFVDDFL 864
b 830 NKKQYQKMGIFQGLCVSVILSSFYANLENA-LQFLAKESMDPEKPEINLLMRLTDDYL 889
Y 865 LVTHPLDQAKTFLSTLVHGVBEYGCINLQKTVNFPVEPOTLGGAAPOYLPDHC----- 919
b 890 LMTTEKNAMVLEIEKLYQLSLGNFFKFMKRLKTNFALNLQIG-----CTNTTQ 939
Y 920 -----LFPKGLLDTQLEVECDYS---GVAOTSIKTSITFOSVFKAGTWNK 967
b 940 DIDSINDDLFWHIGISIDIKTLNIQINIKSGEILCTLVNVMGTNESIILWLKXSLF 999
Y 968 LSVRLKCHGLEFLDLQVNSLQTVNCINIKYIFLQAYRFHAC 1008

```

```

DB 1000 MNNT-----SFYFKSTINTKQFANITLSKLYIAABEKYVAC 1035

RESULT 5
T14891
telomerase (EC 2.7.7.-) catalytic chain p133 - Tetrahymena thermophila
NAlterate names: telomerase reverse transcriptase
CSpecies: Tetrahymena thermophila
CDate: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
CAccession: T14891
RCollins, K.; Gandhi, L.
Proc. Natl. Acad. Sci. U.S.A. 95, 8485-8490, 1998
A>Title: The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex
A.Reference number: 218252; MUID:98337941; PMID:9671704
A.Accession: T14891
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-1117 <COL>
A.Cross-references: EMBL:AF061284; NID:G3335166; PID:G3335167; PIDN:AAC39140.1
C.Genetics:
A.Genetic code: SCS
C.Keywords: nucleotidyltransferase

Query Match      6.8%; Score 401.5; DB 2; Length 1117;
Best Local Similarity 21.2%; Pred. No. 2.7e-21;
Matches 173; Conservative 147; Mismatches 358; Indels 139; Gaps 28;

QY 296 SDSLGSVCCKKPKSPSTLSLSPRONAFQLRPFIETRHLYSRGDQGE----- 344
DB 190 NFNMGKATSNNNQNNANLSNEKQENQYIPEIQRSQIFYCNHMGREPGVFKSSFFNY 249
QY 345 -RLNPSFLLSNLQPNLTGASRLVEIIFLGSRPRT--SGPLCETHLSRRY--WQMRPLF- 398
DB 250 SEIKKGQFQVIOEKLOGRQFINSDKIKDPHOTIKTKLLKEYQ-SKNFSQBEEDLEL 308
QY 399 ---OQLLVNHAEQYVRLRSHCRF-----RTANQOVTALNTSPH 437
DB 309 EFTEKIVQNFHNINFNLYLLKFKCLPENTQSLKSQVKQIVQSENKANQOCSNLFNS--- 365
QY 438 LMDLRLHSSPWQVYGLFRLACLCKVVSASLWGRHNERFFKNLKKFISLGKYGKLSLOE 497
DB 366 ---LYDTEISYKQITNFIROIIONCVPNLLG-KKNFKVLEKLYEFVQMKFENQKVLJ 421
QY 498 LWMKMYVEDCHWLRSSEPKORVPAABH-RLRERILATFLFLMDTVVOLLRSFFYTES 556
DB 422 YICFMDVDFVEWFVDLKNQKFTQKRYISDKRILGLDILVFIINKIVIPVLYNFYITEK 481
QY 557 TFQKRLFFYKRSVMSKQSIGVROHLERVLRLSDEEVRHH---QDTWLAMPICRLRP 613
DB 482 HKGSQYFYRKPFWKLVSKLTI-----VKLEENLEKVEBKLPEDSFQYPOGKURI 535
QY 614 IPKPNGLRPIVNMVSMGTRALGRKQAOHFTORLTKLF--SMLAYERTKH--PHLMGSS 669
DB 536 IPKGSFRPIWTFI-----RKDKQKNILNQLIIMDSQLVFNRLKMLGQKIGYS 586
QY 670 VLGMNDIYRTWRAFLVRALDQTPRMVYFKADVGTGAYDAIPQGLKLEVEVANMIRHS--- 726
DB 587 VFDNKQISXFAQFIEKWNKNGR-PQLYVYVTLDIRKCYDSIDQMKLL-----NFFNQSDLI 641
QY 727 ESTYCIROYAVVRDSDGOVHKSFRQVTTLSLQ-----PYM-----G 765
DB 642 QDTYFINKYLLFOENKRPPLQIOQTNNLSAMEIEBEKINKKPFKMDNINFFYFNLER 701
QY 766 QFLKHLQSDASAL-----RNSVVISQSI SMTESSSSSLFDFFLFLRHSVVKIG 814
DB 702 QIAYSLYDDDDQILQKGFKEIQSDDRFPFVINDKPRCITKDIHNLKHI SQYNVISFN 761
QY 815 DRCTYQCGIPQSGSLSTLLCSLCFGDMENK-----LFAEVORDG---LLRFDVDDFLV 867
DB 762 KVPKQKRGIPQGLNLSVGLCSFYFGKLEBEYTFQFLKNAEQVNGSINLLMRLTDDYL 821

```



788 STSMNESSSLPDLFLHFLHRSVWVKGIDRCYTCQCGIPQSSSLTLCSLCFGDMENKLP 847  
b 409 EMS-----SSTLLSVIRTLSTLYYKRGTSWRITKGVPOQHPISSNLAHMYLNNFEQKYW 465  
y 848 ABVORDGLL--RFVDDFLVTPHLDQAKTFLSTLVHGVPEYGCMLNLOKTVVNFVPEPG 905  
b 466 SNEKEDSRVFCRYEDDDIFITENSLEROMKELSTGNTHFTANPKK-----515  
y 906 TLGGAAPVLPAC-----LFPWCGLLLTQTLFV 936  
b 516 -----FKKSRCCGASQVLCWGVKLDQFQSGNCF 543  
RESULT 8  
95119  
ransposase, uncharacterized, truncation [imported] - Streptococcus pneumoniae (strain R  
;Species: Streptococcus pneumoniae  
;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
;Accession: E98119  
;Host: Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; E  
; R.; Lesland, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
; P.; Sun, P.M.; Winkler, M.E.  
; Bacteriol. 183, 5709-5717, 2001  
;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
;Reference number: A97872; MUID:21429245; PMID:11544234  
;Accession: E98119  
;Status: preliminary  
;Molecule type: DNA  
;Residues: 1-330 <KUR>  
;Cross-references: GB:AE007317; PIDN:AA00786.1; PID:g15459687; GSPDB:GN00174  
;Genetics:  
;Gene: transposase H  
Query Match 2.2%; Score 131.5; DB 2; Length 330;  
Best Local Similarity 19.4%; Pred. No. 0.057; Mismatches 80; Indels 131; Gaps 8;  
Matches 59; Conservative 34;  
607 PICRLFIKPKGLRPIVNMYSMGTALGRKKAQHFTQRLKTLFSLMNYERTKTHLM 666  
b 3 PICEPHSDTSYCFRP--NRSC-----KAIMKLEY-----32  
y 667 GSSVLGNDIYRTWAFVLRVLRALDQTPRMVYKADVTGAYDAIPQGLVEVVMNTRHS 726  
b 33 -----LNDGYE-----MTVIDLEKFFDTVPQDLMSLVHNIIDG 68  
y 727 ESTYCIROYAVVRDQGVHKSFRQVTLSDLPYMGQFLKHLQDSASALRNSVIE 786  
b 69 DTESLRKYL-----78  
y 787 QSISMNESSSLPDLFLHFLHRSVWVKGIDRCYTCQCGIPQSSSLTLCSLCFGDMENKL 846  
b 79 -----HSGVIINGCRYKTLVGTVPQGNLSPLLSNI-----MLNEL 113  
y 847 FAEVQRDGL--LLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGCMLNLOKTVVNFVPEP 904  
b 114 DKEELKGLRFRVYADDVCVTVVSEAAKRMYSVSRFIEKGLGLKVNMTKTLTRPREL 173  
y 905 GTLG 908  
b 174 KYLG 177  
RESULT 9  
38487  
;astin - human  
;Species: Homo sapiens (man)  
;Date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 05-Nov-1999  
;Accession: I38487  
;Fukuda, M.N.; Sato, T.; Nakayama, J.; Klier, G.; Mikami, M.; Aoki, D.; Nozawa, S.  
;enes Dev. 9, 1199-1210, 1995  
;Title: Trophinin and tastin, a novel cell adhesion molecule complex with potential inv

A;Reference number: I38487; MUID:95278733; PMID:7758945  
A;Accession: I38487  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-778 <RES>  
A;Cross-references: EMBL:U04810; NID:g905355; PIDN:AAA79333.1; PID:g905356  
Query Match 2.2%; Score 131; DB 2; Length 778;  
Best Local Similarity 22.4%; Pred. No. 0.22; Mismatches 167; Indels 160; Gaps 26;  
Matches 107; Conservative 44;  
2 TRAPRCFAVRSLLRSRYREVMPLATFVRRLGPEGRRLVQGPDKIYRILVA-----QC 54  
b 358 TPEPSTPRVQ-----AQWLRGVSPQS-----CSEDPALPWEQVAVRLFQDESC 401  
y 55 LVCWHSQSP---PADLSFHOVSILKELVAR---VVORLCERNERNVLAFGFELN 108  
b 402 IRSLEGSGKPPVATPSGPHSNRTFSQEVKIQIGLQLLQKQEVGLVGGQCVPLN---458  
y 109 GGP--PMATSSVRSYLPNTVIETLRVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSC 166  
b 459 GGSLLDMVELQPLLTEISRTLNATEHNSGT-----SHLPGLLKHSL-----PKFC 504  
y 167 AYQVCGSLVQICATTDINP-----SYSASRPTRPVGRNFTN-L 205  
b 505 LPFECGEP--QPCPPAEPGPPEAFCKSEPEIPEPSLQEQLEVPEYPPAEP--RPLESC 560  
y 206 RFLQIKSSRQE-----APKPL-----ALPSRGTKRHLSTLTSVPSAKK 246  
b 561 RSEPEIPESRQEQLEVPEPCPAERPLSEYCYRIEPIESSRQQL-----EYPE---612  
y 247 ARCVPVPRVEGPHRQVLPTPSGKSWVPSPARSPVPTAKDLSSKGKVDLSLGSV---304  
b 613 ----PCPPAEPGP---LQPSTQCGSGPPGPC-----PRVELGASEPCTLEHRSLESSLP 660  
y 305 CCKH-KQSSTSL-----SPPRQNAFLRPFIEHFLYSRGSGGGERLNPSLLSN 354  
b 661 CCSQWAGATTSIFSSOHLPCASPPICLSQSLAP-----PAGQAG-----LSN 703  
y 355 LQPNLTGARRIV-----EIIFLGSRPRTSGP--LCRTHLSRRYVQ 393  
b 704 LAPRTIALRESLKLTAIHCHEARLDDCAFTSRASPSGTRVCTNPVATLLEWQ 761  
RESULT 10  
T24227  
hypothetical protein R166.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T24227  
R;Matthews, P.  
submitted to the EMBL Data Library, August 1995  
A;Reference number: Z19859  
A;Accession: T24227  
A;Status: preliminary; translated from GB/EMBL/DD8J  
A;Molecule type: DNA  
A;Residues: 1-502 <WIL>  
A;Cross-references: EMBL:Z50795; PIDN:CAA90562.1; GSPDB:GN00020; CESP:R166.1  
C;Genetics:  
A;Gene: CESP:R166.1  
A;Map position: 2  
A;Introns: 32/1; 164/1; 306/1; 351/2; 480/3  
Query Match 2.2%; Score 129; DB 2; Length 502;  
Best Local Similarity 22.0%; Pred. No. 0.16; Mismatches 159; Indels 138; Gaps 22;  
Matches 100; Conservative 57;  
141 LSRVGDLLVYLLAHCALYLLVPSQVCGSPLYQICATTD----IMPSVSASVPT 196  
b 70 LAEYSDQTAFLA--ALQOIGPP-----PPLNTYPTAGTFMALLPGLIAATSP--117  
y 197 PVGRNFTNRLFLOQIKSSROEAPKPLALPSRGTKRHLSTLTSVPSAKKARCYVPRVE 256

118 -----KFPRLRLSLSSVAEKVSTPSEAAATSSSSPSLNLTSSNSVPLAKFPSTPLE 173  
257 E-----GPHRQVLPSPGSKSWVPS-----PARSEVPTAEKDLSSK-----G 293  
174 SLASDQSSSTSSVRSVLPSTSSNTSHPELPAGILPATTTNNVAAVPPSPSRATANVFGS 233  
294 KVSDDLSSGSV-CKKH--KPSSTSLSP-----RQNAFLRPPTETREHFLYS 338  
234 NSIGLNFSGAASVTRHLVWPSSSTIQOFTSGRSSSTGQKEGSSSPFLGVG--YS 290  
339 RGDQERLNPFLSLNQLPNLTGARRLVEIFLGSRPRTSGPLCRTHLSRRYQWRPLF 398  
291 QPYGNDVFS--LQDFDEN-----NPSLTENPTLSAQISRL----- 324  
339 QOLLVNHAEQVY-----RLLRSHCRFTANQOVTDALNTSP--PHLMDLLRLHS 446  
325 -----AECALAAKSNLPLPLPVNKRR--VSKEVIELLKCFATPSIHAFRKYS 374  
447 SPWQVYGF-----RACLCKVVSASLWGTNRNRRFFKNLKKFIS 486  
375 A--IYGRFDTKRPHKVLTLHTTVNEAAQLCLLVPSLL--TRDE--LPLARQIVK 427  
487 LKYGKLSQLQELMMKMKVEDCHWLRSPFGKDRVP 520  
428 DAGNYAKSKR--PCDPAHLSPISFGNSPPP 459  
RESULT 11  
32730  
VA-directed DNA polymerase homolog T914.6 - Arabidopsis thaliana  
Alternate names: hypothetical protein At2g28980  
Species: Arabidopsis thaliana (mouse-ear cress)  
Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001  
Accession: T02730; C84691  
Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Roun-  
abmitted to the EMBL Data Library, August 1998  
Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.  
Reference number: Z14710  
Accession: T02730  
Status: translated from GB/EMBL/DBJ  
Molecule type: DNA  
Residues: 1-1529 <ROU>  
Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461840  
Experimental source: cultivar Columbia  
Lib, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.  
iss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
ature 402, 761-768, 1999  
Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
Reference number: A84420; MUID:20083487; PMID:10617197  
Accession: C84691  
Status: Preliminary  
Molecule type: DNA  
Residues: 1-1529 <STO>  
Cross-references: GB:AE002093; NID:g3461840; PIDN:AAC33226.1; GSPDB:GN00139  
Gene: At2g28980; T914.6  
Map position: 2  
Introns: 428/1; 532/3  
Query Match 2.1k; Score 125; DB 2; Length 1529;  
Best Local Similarity 19.2k; Pred No. 1.7; 341; Indels 394; Gaps 58;  
Matches 209; Conservative 147; Mismatches 147; Indels 394; Gaps 58;  
142 SRVGDDLLVYLAHALCALYLVPPSCAYQVCGSPYQICATTDPWPSYASRPTFPGVN 201  
220 SETGVDAWVEY-----KYPWLPFRCSGCKWGHQEVCLTR--PSPNQLSTPT----- 265  
202 FTNLRFLLQIKSSSRQEP-----KPLALPSRGTKRHLSTLTSVPSAKKACYPV--PRV 255  
266 -----ELETEDKTEPPLMKKPEILSKSPSATLTKTLNGDSHTQKV---PMKNPTV 314

RESULT 12  
T30838  
cytoplasmic dynein heavy chain - Paramesium tetraurelia  
C;Species: Paramesium tetraurelia

256 -----BEGPHRQVLPSPGSKSWVPSPAR-----SP-----EVPTA 285  
315 LONKGEVAEENENKDEGPHWTVSPAKVGRQSPQOEVVNVNHPASRQFVLAVEEDTA 374  
286 EKDLSSGKGVSDLSLSSGVCCKHKPSSSTSLSPPRQNAFQLRPFT-----ETRHFLYSRGD 341  
375 DTNGHEEGEIVLSQ-----QLNDNSDSV-----GVQGTPEYIPRISKSHQKVVKSSA 422  
342 QGB-----RLNPSFLLSNL-----OPNLGARRLVEI 368  
423 NQNTKGRIVWVRDNRALTVPFKSSOMITCSILLEGKEBEFFCSFIYASNFVEER---I 479  
369 IFLGSRPRTSGPLCRTHLSRRYQWRPLFQOLLVNHAEQVVRLLRSHCRFTANQOVT 428  
480 LWEDIRSHSDPLIR-----NKPWILCGDFNEILGEGHSNY----- 516  
429 DALNTGPPHMDLLRLHSPWQVYGFRLACLCVKVVSASLWGTNRNRRFFKNLKKFISLG 488  
517 DNSPYTPPGWRD-----FOEIGRL-----MLEAAATGGR--KPF--KFVNV- 553  
489 KYGKLSIQELMMKMKVEDCHWLRSP-----GKORVPAE 523  
554 -----LTKLQPLFPVESHNAASSAPLYVSTGALYRFSKKLTKLPHRLGKELGDL 607  
524 HRLRE-RIL-----ATFLEWLMDTYVQLLRSFFYIT-----ESTFQK----- 560  
608 KRTREAHILLCKEQATTLANPSQETIABELKAYTDWTHLSELEGFLKQSKLHWMVGD 667  
561 -NRLFPYKSVMSKQSIGVROHLERLREL-----SQBEVRHODTWLAMPICR 610  
668 GNNSYFHKAQVRKM-----RNSIREIRGNPAETLOTSEEIKGEARFFN----- 712  
611 LRPIKPNQ-----LRPIVNMVSYNGTRALGRKQAOHQFTQRLKTLFSLMNIYRTK 661  
713 -EFLNQSGDFHGISVEDLRLNLSYRCSVTQDNILTRVTGEEIQ--KYLAMPN--NK 766  
662 HPHLMGSSVGNNDIYR-TWRA-----FVLVRALDQTPRMVYFVADVTGADAIPOGKL 716  
767 SPGPDGYT-----SEFFKATWSLTGPDFAAIQS-----FFVKG-----FLPKGLNA 808  
717 EVVANMIRHSES-----TYCIRQYAVVRDSQGVHKSFRROVTTLSDLQPYMGQF 767  
809 TILALIPKDEAIEMKDVPISCCNVLYKVISKILAN-----RLKLLPSFILQNSAF 862  
768 LKHLDQSDASALNSVW-----IEQISMNNESSSLFD-----FFLHFL----- 806  
863 VKERLLMENVLLATELVXDYHKESVTPRCAMKIDISKAFDSVQVQFLNTLEALNFPETP 922  
807 RH-----SVVKIGDRC--YTQCGIPQSSSLSTLLCSLCFGDMENKL-FASVQR 852  
923 RHWIKLCISTATFSVQVNGELAGFFGSSRGLRQGCALSPYLFVICMNVLSHMDAAVHR 982  
853 D-----GLL-LRFVDDPFL-VPHLDOAKTFLSTLVHGVPEYGCMLNLOK----- 895  
983 NIGYHPKCEKIGLTHLCFADDLVMPVDGHWISGVINVFPKFAGRSGLQISLEKSTIYL 1042  
896 -----TVNVPFVPGTLGGAAAPQLPAHCLFPWCGLLLDTQTLEVPDCYSGVAQT 945  
1043 AGVSA9DRVQTLSSFPFANG-----QLPVRYL-----GLPLITKQMTT-ADYSPLIE- 1088  
946 SIKT-----SUTFQSVFKAGTKMKNLLSVLRKCHGLFDLDQVNSLQTVGINIYKIFLQA 1002  
1089 AVKTKLSSWTARSLSYAGR-----LALLNSVIVSIAN-FWMSA 1125  
1003 YRFHA-CVIQL 1012  
1126 YRLPAGCIREI 1136



Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 02-Mar-2001  
Accession: T30838; PC4341  
Kandl, K.A.; Forney, J.D.; Asai, D.J.  
Submitted to the EMBL Data Library, January 1995  
Description: The dynein genes of Paramecium: the differential expression of axonemal a  
Reference number: Z20502  
Accession: T30838  
Status: Preliminary; translated from GB/EMBL/DBJ  
Molecule type: DNA  
Residues: 1-4540 <KAN>  
Cross-references: EMBL:U20449; NID:G987228; PID:G987229; PIDN:AAA75445.1  
Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.  
Cell Sci. 107, 839-847, 1994  
Title: The dynein genes of Paramecium tetraurelia: Sequences adjacent to the catalytic  
Reference number: PC4340; MUID:94334383; PMID:8056840  
Accession: PC4341  
Molecule type: mRNA  
Residues: 1831-2029 <ASA>  
Genetic code: SGC5  
Introns: 57/3  
Superfamily: dynein heavy chain, cytosolic  
Query Match 2.1%; Score 124; DB 2; Length 4540;  
Best Local Similarity 18.5%; Pred. No. 10;  
Matches 127; Conservative 111; Mismatches 214; Indels 234; Gaps 33;  
Y 352 LSNLPNLTGARRLVEIIFLGSRRPTS-----GPLCRTHRLSRRY 391  
b 356 LSNIGETIIPSL-QLAESRSRELTNEMIKYKGFQQLHKYVDFKGLIKIQTIPSQW 414  
Y 392 WQWRPLFOQLLYN---HABCQYVLLRSHCRFTANQQVTDALNTSPPLMDLLRLHSSP 448  
b 415 DEEYKIFKQSIKVSQVHVKDQYGF--SHIKLOXQIQ-----HIQRLREHNEH- 460  
Y 449 WQVYGLRACLKCVVSASLWGRHNERP-----FKNLKFIKSLGKYGKLSLQ 496  
b 461 -----LKEVIEQIIQNDQOEKENVQFATQETQAYDIFKQVVEF-DLSRDGDEQFF 513  
Y 497 ELM--WKMVVECHW-----LRSSPGQVFAAHRERILATF--LF-----WIMDT 541  
b 514 RALKQVEIAESVEATITNLRDLSGS---ASSAKEMFRILAKFNKLFSPRIKGAIQE 569  
Y 542 YVQVLLRSFFYTES-----TFQK---NRLFYRK-----SVMSKLSQISQVROHLE 584  
b 570 YQSOLLKTVHDKIQSLQNKFKETQYKQNSRLASARDIPLTSGFVWSKQLQIRLQKYM 629  
Y 585 RV-----LRLELQ-----EEVRHQDTWLAMPICRLRF 613  
b 630 KYEQILGPQWAEITGKCKKENGTEFERILDSPGPALEDWQKEINHNN-----KA 678  
Y 614 IPKPNGLRPIVNNYSVMGTRALGRKQKQHTQRLKTLFSLMANTYRTKQHPHMGSSVLGM 673  
b 679 VSQNEKLFVWTRRRGLEIRVNYEKKLSQLF-KEVRNLSNM---KTKVPYS-SHIA--- 730  
Y 674 NDIYRTWRALVLRALDQTPMYFVKADVTCAVDALPQGLVEVWAMHSESTYCIIR 733  
b 731 NDAKASY-FPAL---SLQESLHYI---CITSQLNA-KSAKLIV----- 765  
Y 734 QVAVRRDSQGVHKSFRFRQVTTLSLQFYMGQFLKHI----- 771  
b 766 --AALARKEVQLQIGQGFNYLWTHKTLQLPYVKFTDKVFELEQAVGNLNERIGQIBSLCE 823  
Y 772 -----QDSALRNSVVIOSISWNSSSLPFFLHFLRHSVVKIGRBYTCOCGIP 825  
b 824 AMKTCVPVDSLADKLDIOBVIDSLCFNFSN-----LHIWQIDIX----- 864  
Y 826 QGSSSLTLLCSLCFGDMENKL-----FAEVQRDGLL-----LRFVDDFLLVTPHLDQA 873  
b 865 ---QIESILCDRTVQMKELANQFINYQIKIQERGLVNTVTVHELKLDQIIYDPPVEYA 921  
Y 874 KTLSTLVHGVPEYGCMMINLQTVVN 899

DB 922 KYFWFQFQHKM--IGQICSLPRILVAN 945  
RESULT 13  
E71436  
Hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: Columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 11-Jan-2002  
C:Accession: E71436  
R.Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D  
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; G  
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomene  
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: E71436  
A:Molecule type: DNA  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Residues: 1-2088 <BEV>  
A:Cross-references: GS:297342; NID:G2245031; PID:G2245044  
C:Genetics:  
A:Map position: 4COP9-4G3845  
Query Match 2.1%; Score 121; DB 2; Length 2088;  
Best Local Similarity 19.9%; Pred. No. 5.3;  
Matches 151; Conservative 82; Mismatches 251; Indels 276; Gaps 37;  
Y 88 LCERNERNVLAQFELLNEA---RGPPMAFTSSVRSYLPNTV----- 127  
b 1293 LSEKRRHIVETGLTLTQASVPREYWPYAFAAV--YLINRMPTVLSMESFPQKLF 1350  
Y 128 ---LETURVSGA---WM-----LLSRVGDLLVLLAHCALY--- 159  
b 1351 KPNYERLURVFCGLCPWLRPYTNKLEERSRCVFLCYSLTQYALCFVVEKRLYTSRH 1410  
Y 160 ---LLVP-----PSCAYQVCGSPLY-QICATTDIWPVSASVYRTPRVGRNT 203  
b 1411 VVDEASFPFSSNLTSQNSLPTVTFEQSSPLVTPILSSSSVLPSCLS--PCTVLHQOQF 1468  
Y 204 NLRFLQIKSSROEAPKPLALPSRGTKRHLSTSTSVPS-----A 244  
b 1469 PVTTPNSFHSQPTTSAPLS-PHRTMDTFQVQPTAPNENGPEPEAOSPIGILSNPT 1527  
Y 245 KKARCYVPVVEEGPHRQVLPSPGSKSWVSPSPARSPEVPTAEKXLSKGVSDLSLSGV 304  
b 1528 HEATIGLPNPNRNPTEIEPT-----PAPHPKVPKTTTTTTTNRRTVSDAS---- 1575  
Y 305 CCKHKPSTSLSPRONAFQLPPTETRHFLYSGDGO-ERLNPFSFILSNLPNLTGAR 363  
b 1576 ---HQPTA-----POON-----QHNMKTRAKNNIKKPNTKFSLTATLPNRS--- 1613  
Y 364 RLVEILFSGSRPRTSGPLCRTHLSRRYW-----MRPLFQQLIVNEA 406  
b 1614 -----PSEPTNTVQALKKKWRPAMSDDEFDAQORHNTWDIVPHESQLLVG--- 1658  
Y 407 ECQVRL-----RSHCRFTANQQ---VTDALNTSPHMLMDLRL-----HSSP 448  
b 1659 -CKWYFKLKYLPNGAIDKYKARLVAKGNQYGVDAETFSPIVKSTTIRLVDVAVKDD 1717  
Y 449 WQVYGFLEACICKTVSASLWGRHNERFFKNLKKISLGLKYGKLSLQELMWKMKVEDCH 508  
b 1718 WEIKQL-----DVNNAFLOGTITE----- 1738  
Y 509 WLRSSPG---KDRVPAAEHRLRERILATFLFLMDTVVQLLSFFYI-TESTFQKRLF 564  
b 1739 YMAQPPGFIDKOR-PTHVCRLKAIYG--LKQAPRAWYMLKQHLFNIGVNSLSASL 1795  
Y 565 PYKSVSKQJSGVGRHRLRERLRELSQEVVHQDTWLAMPICRLRFPKENGRLPIV 624

```
1796 IY-----WSDKSSID-----AVLTS:AE-----RF-----SIKDF 1821
625 NMSVSMGTRALGRKQKQHTQR--LKLTFSLMNYERTK-----HPHLMGSSVLGN 674
1822 DLHYFLGIEAT-RTKQGLHLMQRYKNDLLAKHNMAKAPVLTPLTSPKLTUHGKTIN 1880
675 DI--YRTWRAPVLVRALDQTPRMVFVKADVTGAYDAIPQ 712
1881 DASEYRS-----VGSQLYLATRPDIAYVNRLSQ 1911

RESULT 14
10740
Species: Caenorhabditis elegans
Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
Accession: T20740; T23800
Submitted to the EMBL Data Library, December 1995
Reference number: Z19800
Accession: T20740
Status: Preliminary; translated from GB/EMBL/DBDJ
Molecule type: DNA
Residues: 1-1646 <W1>
Cross-references: EMBL:Z68297; PIDN:CAA92597.1; GSPDB:GN00022; CESP:Flila10.4
Experimental source: clone Flila10
Steward, C.
Submitted to the EMBL Data Library, January 1996
Reference number: Z19800
Accession: T23800
Status: Preliminary; translated from GB/EMBL/DBDJ
Molecule type: DNA
Residues: 1-1646 <W12>
Cross-references: EMBL:Z68507; PIDN:CAA92830.1; GSPDB:GN00022; CESP:Flila10.4
Experimental source: clone M18
Genetics:
Gene: CESP:Flila10.4
Map position: 4
Introns: 102/3; 146/3; 265/3; 343/3; 424/1; 460/2; 736/2; 796/3; 891/2; 947/3; 1331/3;

Query Match 2.08; Score 120.5; DB 2; Length 1646;
Best Local Similarity 19.38; Pred No 4.1;
Matches 244; Conservative 159; Mismatches 412; Indels 449; Gaps 64;

49 TLVAQCLVCMHWGSGPP-----PADLS-----PHQVS----- 75
410 TVQQCLVCTFSSDQENEDRSQEDGGPGFLSKGLWPVVEHLTSKTKILLDSLRMDA 469
76 -----SLKELYARVVQRLC-----ERNERNVLAFGPELLNEARGGPPMAF 115
470 VAIEGVYLRCCVACDMTQAVTAIDKLCULDENSEAGSGETKLEIAKAYANSQPSI 529
116 TSSVRSYLPNTVITLRSVGAWMLLS-----RVGDDL-----LVYLAHCA 160
530 LAAIGSLLAASTDEI--VSPQLCCLSTLISACRGVADADLHRSVYVLAIVMSLSPSYL 587
161 -----LVPPSCAYQVCGSPLYQCATTDIWP-----VSASYRTPRPVGRNFTNLRFLQIK 212
588 NQFAGIAPPSPANK-RDAPVSEQVDEAWPSTQVTASGPP----- 628
213 SSSROEAPKPLALPSRGTGRHLSLTSTSVPSAKKACYPVPERVEGPHRQVL--PTPSGK 270
629 -----CPCPVSTELWNKQVL-LTSKNLQAARTFIATITIKELNNLWYLCMATCEHL 681
271 SWVSPARSPEVTPAEKD-----LSSKGKVDLSLGSVCKKXPSSTLSLSPR 320
682 SWLLA-MRPTQVQOFERETDRDHSNPTVVTNAAISDIAMLSLMDKVAPAIALPD--- 737
321 QNAQLAPFIEIETHFLYSRGD-----COE-RLNP-----SFLSNLPNLTGAR 363
738 -----KSFLUSDALIRLDESIAVAATGRESLFLAVLYKVCVSLMLRQ----- 784
364 RLVEIIFLGRPRPTSGPLCTRHLRRY-----QMRPLFQQ 400
```

```
Db 785 -----VFQKAAAHFKVCNHTSVSMRDWAVALTSLAKHAVKSKTSMDAKSQEMIISS 839
Qy 401 LL-----VNHAEQYVLLRLSHCRFTANQQVTDALNTSPHMLDL----- 441
Db 840 LLALCSIPHIQ--VRRQLDCVMSLMQTDGAFLSTSWPNVIQIISAIIDSDTECELSL 896
Qy 442 -----LRLHSSPMQVYGFLEA-----CLCKVVSASLWGTNRNRRFFKMLKFLISLGY 490
Db 897 VRQGYLGLRLVSS-----DFLOSIFDFCISGLVEAI-----SRYSKONTDONI- 939
Qy 491 GKLSLQELMWKM-----KVEDCHWLRSRPGKDRVPAAEHLRLRILATFLFLWMDTVV 544
Db 940 -SLSALTLLWTISDFYRKE-----SVGND-----ASEAVMVV----- 972
Qy 545 QLLRSFYITESTFOKNRLFFYRKSV-WSKQSGVROHLERVLRELBSQBEVRHQDTW 603
Db 973 ---LVTCLSESCVDSR--FAVRKSAQCTLLQTVTAHGAHLR-----SAAWHSVIW 1017
Qy 604 -LAMP1-----CRLRFIPKPNGLRPIVNNMSYMGTRALGRRKQAOHFTQRLKTLPSML 655
Db 1018 QIMIPLLDKVRSQTRCASTEKSG--ELIMHSDTE---QKWTETCIHTISAIKIP 1071
Qy 656 NYERTKHPLMGSSVLGMNDIYRTWRAPV--LRVRALDQTPRMVFVKADVTGAYDAIPQ 713
Db 1072 NSOR-----KSLIALNDFGAVWEAFGLVLDWAACYENAE--SLSAIRSQEVLG 1120
Qy 714 KLVEVVANMIRHSESTYCTIRQAVVRRDSQGVHKSFRQVTTLSDLQPMGQ----- 766
Db 1121 KISSQTLNVNSHEKN-----GSDS-----TIDAITPELPQAWVESW 1158
Qy 767 -----FLKHLQSDSALRNSVIE-QSIS-----MNESSSLDFDFLFLRHSVVKIGDR 816
Db 1159 KWLRIISRGLAQGCAAMANSYNADSKSISSTPRMNSSSSL-----ASLAPG 1206
Qy 817 CYTCQOGIPQSSSLTLLCSLCFGDMENKLFPAEVORDGLLRFDVDFLLVTPHLDQAKTF 876
Db 1207 IY-----VPGPSHLTAIL--HVFPFLFDKAKSIT--VDDL----- 1238
Qy 877 LSTLVHGPVEYGCINLOKTVNVFVPGTLGGAPYOLPA---HCLFPMGGLLDLTQTL 933
Db 1239 -----KYESLPAVLESMMNVIP-----SEQAPFVLPSATHTLPTQALLAEVAKIV 1285
Qy 934 EYFCDSYGAQTSIKTSLTFQSVFKAGTKMKNKLLSVLRKCHGLFDLQVNSLQTVGIN 993
Db 1286 FVECTLSG--TSLRAAIPDQ-----IRLLKFASTATORIS 1319
Qy 994 IYKIFL--LQAYRFHACYQLPFDQ-RVRKNLTFLGLIISQ-----ASCCVAILKVNPG 1046
Db 1320 PNKAPGGQKSTRDALTIVFSEYSIRIAIEFTTS--TSQNEVANSLIAIDIIKFLG 1377
Qy 1047 MTLKASGSPFPPAAHWLCYQAFLLKLAHSHVYKCLGLPLRTAOKLCKRLPEATMTILK 1106
Db 1378 -----EPLYMKYTCISASTWLAATS-----LMSVLRTSIP-YARQNPFVEFRGLWS 1422
Qy 1107 AAAD 1110
Db 1423 ATCD 1426
```

## RESULT 15

```
Tl1216
reverse transcriptase homolog ORF544 - red alga (Porphyra purpurea) mitochondrion
C;Species: Mitochondrion Porphyra purpurea
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: Tl1216
R;Burger, G.; Saint-Louis, D.; Gray, M.W.; Lang, B.F.
submitted to the EMBL Data Library, December 1998
A;Description: Complete sequence of the mitochondrial DNA of the red alga, Porphyra pu
A;Reference number: Z17255
A;Accession: Tl1216
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
```

```
;Residues: 1-544 <BUR>
;Cross-References: EMBL:AF114794; NID:G4106927; PID:G4106928; PIDN:AAD03095.1
;Genetics:
;Genome: mitochondrion
;Note: orf544
;Keywords: mitochondrion

Query Match      2.0%; Score 116.5; DB 2; Length 544;
Best Local Similarity 21.7%; Pred. No. 1.6;
Matches 75; Conservative 41; Mismatches 94; Indels 135; Gaps 16;

Y 573 KLGSGVGRHLEFVRRLRELSOBEVRHHQDTWLAMPICKRLRFPKPKNGRLRFINVMSYSGMT 632
b 70 KISSITPDERMELVRNIHVQNSDK-----ILRVT-IPKPNG-----SV 107
Y 633 RALGRKQAQHTQRLKTLFSLNVERTKHPHLMGSSVLGMNDIYRTWR-----AFVLRVR 688
b 108 RNLGIPTTRAKQCLVKFALEPQVEAFEPNSYG-----FRPGSANDARKAIVK 158
Y 689 ALDOTPRYFVKADYTGAYDAIPQGLVEVVANNIRHSESTYCIROYAVVRDSQGQVHK 748
b 159 CLQRPK-HILDADINGCFDNIHDKLEKL-----NIFPL----- 193
Y 749 SPFRQVTT-----LSDLQPYMGQFLKHLQSDASALENSVVIQISMSNESSSLFDF 802
b 194 -FRNQINTWLKVGVLSDFG-----NKTETIPOS----- 221
Y 803 LHFLRHSVVKIGRCYTCQCGIPQGSLSLTLLCSLCFGDMENKLFQVQDGLLLRFVDD 862
b 222 -----GTPQGGTISPLLANIALHGIE-KLVS--KRGVYLIRYADD 258
Y 863 FLIV---TEHLDQAK-----TFLSTLVHGVPYEGCMINLQKTVVNF 900
b 259 FLVLCNEEKELLEAKQKIEVFLKTI-----GLQFSAEKTKITY 296
```

Search completed: February 4, 2004, 14:12:28  
Job time : 32 secs

sult No.	Score	Query		ID	Description
		Match	Length		
1	5901	100.0	1122	1	TERT MOUSE
2	3505	59.4	1132	1	TERT HUMAN
3	621.5	10.5	988	1	TERT SCHPO
4	426.5	7.2	1132	1	TERT OXYRTR
5	410.5	7.0	1031	1	TERT EUPAE
6	401.5	6.8	1117	1	TERT TERTH
7	341	5.8	884	1	TERT YEAST
8	131	2.2	778	1	TAST HUMAN
9	124	2.1	454.0	1	DYHC PARTE
10	119	2.0	873	1	GLND VREU
11	116	2.0	881	1	GLND VIBCH
12	116	2.0	968	1	SNXD HUMAN
13	114.5	1.9	1199	1	Y173 HUMAN
14	113.5	1.9	653	1	DEND RAT
15	111.5	1.9	1415	1	ICP4 HSVMG
16	110	1.9	812	1	EBN3 EBV
17	108	1.8	880	1	ARG2 YEAST
18	108	1.8	983	1	4ET MOUSE
19	106	1.8	958	1	MSH4 MOUSE
20	105.5	1.8	901	1	R813 ARATH
21	105	1.8	1473	1	NALI HUMAN
22	105	1.8	1493	1	M3K1 RAT
23	104.5	1.8	916	1	RTVK DROFU
24	104.5	1.8	3511	1	MY15 MOUSE
25	103	1.7	864	1	YC18 HUMAN
26	102.5	1.7	788	1	CY14 NEUCR
27	102.5	1.7	2335	1	TOR1 SCHPO
28	102.5	1.7	2482	1	VWF FIG
29	102	1.7	1493	1	M3K1 MOUSE
30	102	1.7	1795	1	ESPT HUMAN
31	102	1.7	1815	1	SHK3 RAT
32	101	1.7	677	1	T2D5 HUMAN
33	101	1.7	994	1	MERK MOUSE

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF015950; AAC51672.1; -  
EMBL; AF018167; AAC51724.1; -  
EMBL; AF128894; AAD30037.1; -  
EMBL; AF128893; AAD30037.1; JOINED.  
EMBL; AY007685; AAG23289.1; -  
PIR; TC3844; TC3844.  
Genew; HGNC:11730; TERT.  
MIM; 167270; -  
GO; GO:0005696; C:telomere; TAS.  
GO; GO:0003721; F:telomeric template RNA reverse transcriptase. . .; TAS.  
GO; GO:0007003; P:telomere binding; TAS.  
InterPro; IPR000477; RVTse.  
InterPro; IPR003545; Telomerase\_RT.  
Pfam; PF00078; Ivt; 1.  
PRINTS; PR01365; TELOMERASERT  
Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
DNA-binding. 516 516 D -> G (IN REF. 2).  
CONFLICT 1132 AA; 126996 MW; 94E35469C4CA33A0 CRC64;  
SEQUENCE 1132 AA; 126996 MW; 94E35469C4CA33A0 CRC64;

Query Match 59.4%; Score 3505; DB 1; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 2.5e-251;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVRSLLRSRYREVPVPLATVRLRLGEGRLVQDPKLYRTLVLAQCLVCMHW 60  
2 MTRAPRCVRSLLRSRYREVPVPLATVRLRLGEGRLVQDPKLYRTLVLAQCLVCMHW 60  
3 GSOPPADLSFHVSSKELVARVQVRLCERNENVLAFGLFELLNEARGGPPMAFTSSVR 120  
4 DARPPPAAPSPFQSCUKELVARVQVRLCERNENVLAFGLFELLNEARGGPPMAFTSSVR 120  
5 SYLPTNTVTLRVSGAMMLLSRYGDDLLVYLLAHCALYLLVPFSCAYQVCGSPLYQICA 180  
6 SYLPTNTVTLRVSGAMMLLSRYGDDLLVYLLAHCALYLLVPFSCAYQVCGSPLYQICA 180  
7 TTDWPSVASYRTRPVGRNFTNRLFOQIKSSROEAPKPLALPSRGTKRLSLTSTS 240  
8 ATQARPPPHAS-GPRRLG-----CERAWNHSVEAGVPLGLPAPGARRGGASRS 231  
9 VPSAKKACVCPVPRVEEGP-----HRQVLPSPGKSM-VPSFAPSPPEVPTAEKDLSSK 292  
10 LPLKPRGGAAPERTPVCGSWAHPGTRGDSGFCVSPAR-----PAEERTSLE 286  
11 GKVDLSLS-GSVCKKHPSTLSLSPRONAFQLRP-FIETHFLYSGDGOERLNPSF 350  
12 CALSGTRHSFVSQHQHAGPPSTSRPPEPMDTPCPVVAETKHFLYSGD-KEQLRPSF 345  
13 LLSLNQPLNLTGARRIVETILFGSPRTSGPLCRHLSRYWOMRPLFOOLLVNHAEQY 410  
14 LLSURPSETGARRIVETILFGSPRTSGPLCRHLSRYWOMRPLFOOLLVNHAEQY 405  
15 VALLRSHCRFTANQOVDAL-----NTSPPLMDLLRLHSSPWQY 452  
16 GVLTKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRRLVQLLRQHSSPWQY 462  
17 GFLRACLCKVYSASLWTRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWRS 512  
18 GFLRACLCKVYSASLWTRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWRS 522  
19 SPGRDVPAAEHLRERILATFLFWMIDTVYVQLLRFFVITESTFOKRLFFYRSVWS 572  
20 SPGVCCVPAEHLRERILATFLFWMIDTVYVQLLRFFVITESTFOKRLFFYRSVWS 582  
21 KQISGVQHLERVLRLSELSEFVHHQDWTWAMPICRLRFPKPNGLRPIVNMVSMGT 632

583 KLSIGIRCHLXRVQLRESEAEVQRHREARFALLTSRLRIPKPDGLRPIVNMVYVGA 642  
533 RALGRKQKQHFQRLKTLFSLNTERKYPHLMGSSVLGNDIYRTWRPFLVRALDQ 692  
643 RTRFRKRAERLTSRKALFSLNTERARPPGLLGASVLGDDIHRWRTFVLRYAQDP 702  
693 TRPMYFKADVTGAYDAIPQKGLVEVANNMHSSTYCIQYAVVRDSDQGVHKSFR 752  
703 PPELYFVKVDVTGAYDTIPQDRLETVIASIILK-PONTYCVREYAVVQKAAGHVRKAFKS 761  
753 QVTTLSLOPYMGQFLKHLQDSALRNSVVIEOSIGNSSSSSLFDFFLHPLRHSVVK 812  
762 HVSTLIDLOPYMRFQFVAHLQET--SPLRDVAVIEQSSLINEASSGLFDFVLFMFCHAVR 819  
813 IGRCVTCOGIPQGSLSLTLCSLCFCGDMENKLFABVQDRCGLLRFLVDLFLVTPHLDQ 872  
820 IRKSVYQCGIPQGSLSLTLCSLCIGDMENKLFAGIRDDGLLRVDDFLVTPHLDQ 879  
873 AKTFLSTLVHGVPEYGCMLNLTQTVNVFVEPTGLGGAAPYQVLPALHCLFPMCGLLDTQT 932  
880 AKTFLSTLVHGVPEYGCMLNLTQTVNVFVEPTGLGGAAPYQVLPALHCLFPMCGLLDTQT 939  
933 LEVPCDYSYGAQTSIKTSITFSQVFKAGTKWENKLSVLRLKCHGLFLDLQVNSLTQVCI 992  
940 LEVQSDYSYGAQTSIKTSITFSQVFKAGTKWENKLSVLRLKCHGLFLDLQVNSLTQVCI 999  
993 NIYKIFLLOAYREHACVIOQLPQDQVRVKNLTFFLGTIISQASCCYAILKVNKPMGLKAS 1052  
1000 NIYKIFLLOAYREHACVIOQLPQDQVRVKNLTFFLGTIISQASCCYAILKVNKPMGLKAS 1059  
1053 GS---PPPEAAHVLQVAFELKLAHSLVYKCLLPLTAQKLLCRKLPKPEATMILKAAA 1109  
1060 GAAGPLPSEAVQMLCHQAFELKLTTRHVTYVPELGLSLRQTLQSLRKLPGTTLTAAEA 1119  
1110 DPALSTDFOTILD 1122  
1120 NPALPSDFKTLID 1132

## RESULT 3

TERT SCHPO  
ID TERT SCHPO STANDARD; PRT; 988 AA.  
AC 013339; 013338;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Telomerase reverse transcriptase (BC 2.7.7.-) (Telomerase catalytic subunit).  
GN TRT1 OR SPBC29A3.14C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=97400623; PubMed=9252327;  
RA Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H.,  
RA Lingner J., Harley C.B., Cech T.R.;  
RT "Telomerase catalytic subunit homologs from fission yeast and human.";  
RL Science 277:955-959(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwalliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis B., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris B., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., A. Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreanc S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., A. Shpakovski G.V., Ussery D., Bazrell B.G., Nurse P.; T "The genome sequence of Schizosaccharomyces pombe."; L Nature 415:871-880(2002).

C -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.

C -!- SUBCELLULAR LOCATION: Nuclear.

C -!- ALTERNATIVE PRODUCTS:

C Event=alternative splicing; Named isoforms=2;

C Name=1;

C IsoId=O13339-1; Sequence=Displayed;

C Name=2;

C IsoId=O13339-2; Sequence=VSP\_006395;

C -!- MISCELLANEOUS: DELETION CAUSES TELOMERE SHORTENING AND SENESCENCE.

C -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.

C TELOMERASE SUBFAMILY

C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

C -----

C EMBL; AF015783; AAC49802.1; -

C EMBL; AF015783; AAC49803.1; -

C EMBL; AL022299; CAAL8391.1; -

C PIR; T03838; T03838.

C GenBank; SPombe; SPBC29A3.14c; -

C InterPro; IPR000477; RVase.

C InterPro; IPR003545; Telomerase\_RT.

C Pfam; PF00078; rvt; 2.

C PRINTS; PR01365; TELOMERASERT.

C Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;

C DNA-binding; Alternative splicing.

C VARSP:IC 524 524 K -> KQ (in isoform 2).

C /FTid=VSP\_006395.

C SEQUENCE 988 AA; 116328 MW; ABZDC7030228F443 CRC64;

Query Match 10.5%; Score 621.5; DB 1; Length 988;

Best Local Similarity 23.6%; Pred. No. 7.1e-38;

Matches 253; Conservative 193; Mismatches 384; Indels 243; Gaps 45;

2y 76 SLVELAVRIVQVLCERN---ERNVLAFFGELLNE-ARGPPMAFTSSVRSYLPNTVIETL 131

3b 83 SQBELIANVVKQMFDSERRNLLMKFSMNHDFRANHVNGVQNDLVSTFFNYLSIL 142

2y 132 RVSGAWMLLSRVGDDLLVYLLAHCALYLLVPFSCAYQVCGSPLYQICATTDIMPSVSAS 191

3b 143 E-SKWQLLEIGSDAMHYLLSKGSIFFALPNDNYLQISGIFLKF----- 187

2y 192 YRTRPVGRNFTNLRFLQIKSSRSRQEAEPKIALPSRGTKRLHLSITSV-----PS 243

3b 188 -----NNV--FEETVSKRKRTTETSITQNKARKEVSNWSISIRSFIFRYS 234

244 AKKARCVPVRBEGPHRQVLPTPSGKSWPSPARSPEVPTAEKDLSSKKGKVDLSLS-G 302

235 YKPKF-----ODLYFNLUH 247

303 SVCCCKHPSSTLSLPPRQ---NAFLQAPPIETRHFLYSGDQBERLNPSPFLSNLQPN 358

248 SICDENTVHWMLQWIPRQGLINAFQVKQL-----HKVPLVSQSTVVP- 292

359 LTGARRLVEIFLGSRPRTSGPLCRTHRLSRVYQWQPLFQQLLVNHAECQVRLLRSHC 418

293 ---KLLKVY---PLIEQTAKRLHRIS-----LSKVYNH-YCPYI----- 325

419 RFRANQOVTDALNTSPHLMDDLRLHSSPQVYGFELACLCKVVSASLWGTGRHNRER- 477

326 --DTHDE-----KILSVSLKPNQVAFRLSILVRVPEKLWG---NORIFE 367

478 --FKNLKKFISLGKYGKLSLOELMWKMKVEDCHWL-----RSSPGKDRVPAAEHRLRERIL 531

368 IILKDLTEFLKLSRYESFSLHYLMSNIKISEIEMVLGKRSNA---KMLSDPFEKKQIF 424

532 ATFLFWMIDTVVQLLRSFFYITESFQKRLFFYKSVMSKLSQSIGVQCHLERVRL--- 588

425 AEFYIWLNSFIPILOSFFYITESDRLNRTVYFRKDINKLL-----CRPFITSMKMEAF 480

589 RELSQEEVRRHQDTW-LAMPICRLRFPKNGRLPIVNMYSYMGTRALGRKQAOAHTQR 647

481 EKINENNVR--MDTQKTTLPVAILPKKNTFLITNLKRELKMGSKNKMVLSTNQT 538

648 LKTLFMSLNVERTKXPHLMG--SSVLGMN-DIYR---TWRAFVLRVALDQTPRMVFKVA 701

539 LRPVASILK-----HLINESGIPFNLEVTKLLTKKDLKLRMPGR--KKYFVRI 589

702 DVTGAYDAIPQCKLVEVVMNIRHSESTYCIROYAVVRDSQGVKHSFRQVTTLSDLQ 761

590 DIKCYDRIKQDLMERIVKXKLKDPK--FVIRKYATIHATSD-RATKNFVSEAFSYDMV 646

762 PY--MGQFLKHLQSDASALRNSVVIQSSIMNESSSLPDFLHFLRHSVVKIGDRCYT 819

647 PFEKVQQLS--MKTSD-----TLFVDFVDYWTYKSSSEIFPKMLKEHLSGHIVKIGNSQYL 699

820 CQCGIPQSSSLTLLCSLCFGDMENKLFAEVQDQG--LLLRFDVDDFLVTPHLDQAKTFLS 878

700 QKVGIPOQSILSSFLCHFYMEDLIDEVLSFKKKGSVLLAVDVFILFVTKKDAKKFLN 759

879 TLVHGVPYEGCMINLQKTVNVPVPEOTLGGAPYQYLPACLPFPWCGLLDQTLEVP-- 936

760 LSLRGFEKHNFSTSEKTVINFENSNGIINNT--FFNESKRMPPFFGFSYNMRSLETLIA 817

937 ---CDYSGYAOTSIK-TSLTFQSVFKAGTKMRNKLKSLVLRKCHGLFLLDQVNSLQTVCI 992

818 CPKIDEALFNSTVELTKHMKGSFFY--KILRSLASFAQ-----VFIDITHNSKFNSSCC 870

993 NIYKI-----FLLOAYRFHACVIOLP-----FDQVRKNLTPFLGIITSSQASCC 1036

871 NIYRLGYSMCVQRAQAYLKBKMDIPIQRMFITDLLNVIGRKWKKLAEILGYTSRRFLSS 930

1037 YAILKVNQPGMT--LKASGSPPEAAHWLCYQAFLLKLAHSAVYKCLLGLPLR 1087

931 AEVKWLFCLGMRDGLKSPFKYHP-----CFEQLIYQFQSLTDLIK-----PLR 973

RESULT 4

IDENT\_OXYTR

ID TERT\_OXYTR STANDARD; PRT; 1132 AA.

AC O76332;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic subunit) (Telomerase subunit P133).

GN TERT.

OS Oxytricha trifallax.



543 VLKVFEDLAIITMRCYFYSTKAKYORIFYYRQKINWMTIRLUSIDDLKQ-NLKQVEK 601

594 EEVRHHODTFLAMPICRLRFIPKPNGLRPINVMSSYMGTRALGRKQAOHFTORLKTFLS 653

602 KEMRIFCESQNPAP-KGLRLIPKGDFTFPIMTFNKRKIPNQ-VGKFGSRWTTNNKILQTAHM 659

654 ML-NYERTKPHLMGSSVLGMNDIYRTWRAFVLIRVRALDQTPRMVFKADVTGAYDAIPQ 712

660 MLRKNLKMKFKGSGFAVFNYYDDIMKRYENFVQKWQIN-SPKLIYFVAMDEIKCYDNVD- 717

713 GKLVEVVANNIRHS---ESTYCTIROVAVVRDRSQGVHKS-FRRQVTTLSLQPYMGQFL 768

718 --CERVNVFLQKSLDMXKSEYFILNTEVLKRNKNIIVERSNFRK---LPKQYFRYKFP 770

769 K-----HLQD--SDASALRNSVIEQISWNSSSSLPFFFLHFLHRHSVVKI 813

771 KIGIDGSSYPTLFEILDEDFNLN-MKRTIYIEQQRKKFPKNLLOFVLKICQNNYVF 829

814 GDRCYTCQGIPOGSSLSLTLLCSLQCGDMENKLPFAEVRD-----GLLIRFVDDFL 864

830 NKQYKQKMGKIPGGLCVSVILSYFYANLEENALQFLAKESWDEKPEINLLMLRTDYL 889

865 LVTPHLDOAKTFSTLIVHGVPBYGCMINLQKTVNFPVPEPTLGGAAAPYQLPAHC----- 919

890 LMTTEKNAMLFIEKLYQLSLGNFFKFMKKLKTNFALNLQKIG-----CTNTTQ 939

920 -----LPPWGGILLDTQTELEVCDYS---GVAQTSIKTSLTFSQFKAGKTWENKL 967

940 DIDINDLPHWIGISIDIKTLNINIQNIKEGILCTLVNMQTNESILWLKKUKSFL 999

968 LSVLRXCHGLFLDLQVNSLQTVNCINIKFILLQAYRFHAC 1008

1000 MNNI-----SPYFKSTINTKQPANITLSKLYIAAAEKYVAC 1035

RESULT 5

ID	TEXT	EUPAE	STANDARD;	PRT;	1031	AA.
AC	O00939;					
DT	30-MAY-2000 (Rel. 39, Created)					
DT	30-MAY-2000 (Rel. 39, Last sequence update)					
DT	30-MAY-2000 (Rel. 39, Last annotation update)					
DE	Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic subunit) (Telomerase subunit P123).					
DE	Euplotes aediculatus.					
OC	Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;					
OC	Euplotida; Euplotidae; Euplotes.					
OX	NCBI_TaxID=5540;					
OX	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=97274210; PubMed=9110970;					
RX	Lingner J., Hughes T.R., Shevchenko A., Mann M., Lundblad V.,					
RA	Cech T.R.;					
RT	"Reverse transcriptase motifs in the catalytic subunit of					
RT	telomerase".					
RT	Science 276:561-567(1997).					
CC	-!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR					
CC	THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT					
CC	ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS					
CC	SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE					
CC	SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.					
CC	-!- SUBCELLULAR LOCATION: Nuclear.					
CC	-!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.					
CC	TELOMERASE SUBFAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by, and for commercial					
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a>					
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
CC	-----					

DR EMBL; U95964; AAC47515.1; -.  
 DR InterPro; IPR000477; RVTse.  
 DR InterPro; IPR003545; Telomerase\_RT.  
 DR Pfam; PF00078; rvt; 1.  
 DR PRINTS; PR01365; TELOMERASERT.  
 CW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
 CW DNA-binding.  
 SQ SEQUENCE 1031 AA; 122562 MW; 57B87A63A1FED60F CRC64;

Query Match 7.0%; Score 410.5; DB 1; Length 1031;  
 Best Local Similarity 20.8%; Pred. No. 3.2e-22;  
 Matches 132; Conservative 128; Mismatches 265; Indels 109; Gaps 20;

2Y 470 TRINERRRFKNLKKFTSLQYKGLSLOELMWKVEDCHWLSSPGKDRVPAAEHLRER 529  
 3Y 376 TGRNRKFKVKKYVELANKHELHKNLLELINTREISMMQVTSKHFYFDFHE-NIY 434  
 2Y 530 ILATFLFWMMDTVVOLLRSFFVITESTQKRLFFYKRSVMSKLSQIGVQHLRVLRL 589  
 2Y 435 VLKLLRWLFEDLVSLIRCFVFTVQKSYKTYRKNIWDMKMSIAD-LKETLA 493  
 2Y 590 ELSQEVRHHQDTW---LAMPICRLRPIPKNGLRPIVMSYMGTRALGRKKAQHTQ 646  
 2Y 494 EVOEKEY---EWWKSLGFAPGKRLIPKTTTPRIMFNKI-----VNSDRKTT 541  
 2Y 647 RLKTLFSLNVERTKPHLM-----GSLVGMNDIYETWFAFLVLRALDQT 693  
 2Y 542 KLTNTKILN-----SHLMKTLKRMFKDFGFAVNYDDVMKKYEEFVCKWKVQVQ- 594  
 2Y 694 PRMYFKADVGTAYDAIPQGLVVEVANNIRSESTYCIQYAVVRDSQGVHKSPRRQ 753  
 2Y 595 PKLFFATMDIEKCYDVNRKLSFLTKTKLLSSDFWIMTAQILKRNKNIVIDSGNFRKK 654  
 2Y 754 VTLLSDLPYWGQFLKHL-----QSDASALNSVJLEQSSIMNESSS 797  
 2Y 655 -----EMKDTFRQFKALLEGQYPTLSVLENEQNDLNA-KTLIVKQKRNFFKDN 708  
 2Y 798 LDFDFLHFLRHSVVKIGDRCTOCQIGPQSSILTLCSLCPGDMENKLFPAEVORD--- 853  
 2Y 709 LLOFVINICQYVNFNGFKYKTKGIPQGLCVSSILSSFYVATLEESSLGLRDESMP 768  
 2Y 854 -----GILLRFVDVDFLLVPHLDQATFTSLVHGVPEYGCMINQKTVNFPVEPTLG 908  
 2Y 769 ENPNVNLMLTDDYLLITQENNAVLFIKILNVSRENGFKFNKGLQTFPLSPSKF- 827  
 2Y 909 GAAPYQLPA-----HCLFPWCGLLDDTOTLE-----VFCDSYGAQTSIK 948  
 2Y 828 --AKYGVDSVEEQNIYQDYC--DWIGISIDMKTLALMPNINLRIEGLICTNLNQTK-- 881  
 2Y 949 TSLTFQSVFKAGTKMRNKLISVLRLK-CHGLFLDLQVNSLQTVCTINIKFILQAYRPHA 1007  
 2Y 882 -----KASMLKCKLSFLMNNITHYFRKTIITTEDFANKTLN--KLFTSGCYKTMQ 930  
 2Y 1008 CVIQLPDQVRKNLNFPLGLISSQASCCVAILK 1041  
 2Y 931 CAKE--YKDFKKNLAMSMDLEVSILYSVTR 962

## RESULT 6

TERT TERTH STANDARD; PRT; 1117 AA.  
 ID TERT TERTH  
 AC 077448;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Telomerase reverse transcriptase (SC 2.7.7.-) (Telomerase catalytic subunit) (Telomerase subunit P133).  
 GN TERT.  
 OS Tetrahymena thermophila.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
 CC Tetrahymenina; Tetrahymena.  
 CX NCBI\_TaxID=5911;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=B7;  
 RX MEDLINE=98337940; PubMed=9671703;  
 RA Bryan T.M., Sperger J.M., Chapman K.B., Cech T.R.;  
 RT "Telomerase reverse transcriptase genes identified in Tetrahymena  
 RT thermophila and Oxytricha trifallax";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8479-8484 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98337941; PubMed=9671704;  
 RA Collins K., Gandhi L.;  
 RT "The reverse transcriptase component of the Tetrahymena telomerase  
 RT ribonucleoprotein complex";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8485-8490 (1998).  
 CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR  
 CC THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT  
 CC ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS  
 CC SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE  
 CC SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.  
 CC TELOMERASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AF062652; AAC39135.1; -.  
 DR EMBL; AF061284; AAC39140.1; -.  
 DR PIR; T14891; T14891.  
 DR InterPro; IPR000477; RVTse.  
 DR InterPro; IPR003545; Telomerase\_RT.  
 DR Pfam; PF00078; rvt; 2.  
 DR PRINTS; PR01365; TELOMERASERT.  
 KW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
 KW DNA-binding.  
 SQ SEQUENCE 1117 AA; 133317 MW; E5AP15E86B0F0CD8 CRC64;

Query Match 6.8%; Score 401.5; DB 1; Length 1117;  
 Best Local Similarity 21.2%; Pred. No. 1.7e-21;  
 Matches 173; Conservative 147; Mismatches 358; Indels 139; Gaps 28;

QY 296 SDSLSSGVCCCHKPSSSTLLSPRONAFQLRPFIETRHFLYSGDQGE----- 344  
 DB 190 NPFNMGKATSNNNQNNANLSNKKQENQIYPIQSQIFYNHMGREPGVPKSPFN 249  
 QY 345 -RLNPSFLSLNLPNLTGARLVEIIFLGSPPRT--SGPLCRTHRLSRRY--WQMRPLF- 398  
 DB 250 SEIKKGFGFKVQIEKLGKQGFINSKDKIPQHPQTIKKTLKKEQ-SKNFSCQBERDLFL 308  
 QY 399 ---QQLLVNHAECQVRLLRSHCRF-----RTANQQVTDALNTSPPH 437  
 DB 309 EFTKIVQNFHNFNINFLKFKCKLPENYQSLKSQVKQIVQSENKANKQSCENLFS--- 355  
 QY 438 LMDLLEHSSPMQVYGFRLACLCKVWSASLWGTGRHNRFRFFKLLKFKISLGYKSLQ 497  
 DB 366 ---LYDTEISYKQITNFLRQIQNCVNPQLLG-KCNFKVLEKLYEFVQMKRFENQKVL 421  
 QY 498 LMKMKVEDCHWLSSPGKDRVPAAEH-RLBERLATFLWMDTVVOLLRSFFVITES 556  
 DB 422 YICFMDVDFVEMFVLDKNQKFTQKRKIYISDKRKLGLDILVIPVIRNFYITEK 481  
 QY 557 TPQKRLFFYKRSVMSKLSQIGVQHLRVLRLSLEVRHH---QDTWLAMPICRLRF 613  
 DB 482 HXEGSQIYFKRPIWKLVSCLTI-----VKLEENLEKVEEKLIPDSFKYPOGKLR 535  
 QY 614 IPKPNGLRPIVMSYMGTRALGRKKAQHTQSLKTLF--SMLNVERTKH--PHLMGSS 669  
 DB 536 IPKGSFRPIMTFL-----BQKQKNIKLNQILMDSQLVFNRLKMLGQKIGYS 586



Y 679 TWRAFLRV--RALDQTPRMVFKADVTGAYDAIPQGLVWVWMMIRHSESTYCIQY- 735  
 b 505 RKEFKQRLKKFNVLPELYFMKFDVKSCYDSIPRMECNRLKDALKNENGFFVRSQYF 564  
 Y 736 -----AVPRDSQGVHKSFRQVTTLSDLQPYMGQFLKHLQDSASALNSVV 784  
 b 565 FNTNTGVKLNVNNAVRPKPYELIDNVTV-----HLSNQDV----- 604  
 Y 785 IEQSISSMNESSSLDFFLHFLRHSVVKIGRCYTCQGG:PGSSSLTLLCSLCFGDM-- 842  
 b 605 -----INVMEIF-----KTALWVEDKCYREDGLFGSSLSAPIVDLVYDILLE 650  
 Y 843 -ENKLPFAEVRDGLLRVDFLLVTHPLDQAKTFLSTLVHGVPEYCCMNLQXTVNF 901  
 b 651 FYSEFKAPSODTLILKLADFLIISTDQOQVINKKLMGFGQKYNANRDKIL----- 706  
 Y 902 VEPGTGLGAAPQLPAHCLFWCGLLDTQLEVFCDYSGVYQTSIKTSLTFQSVFKA 961  
 b 707 -----AVSQSDDDTVIQFCAMGHFVKEVWKHSS-----TMNFIHRSKSK 750  
 Y 962 TWRNKLVLRLKCHGLFLDQVNSLQVNCINI-----YK-----IFLQAYR 1004  
 b 751 GFERSIALFNTIRISYKIDTNLSTNTVLMQIDHVVKNISECYKSAFKDLSINTQNM 810  
 Y 1005 RHACVQLPFQDORV 1018  
 b 811 FHS-----FLORI 818

RESULT 8  
 AST HUMAN STANDARD; PRT; 778 AA.  
 C Q12615;  
 T 30-MAY-2000 (Rel. 39, Created)  
 T 30-MAY-2000 (Rel. 39, Last sequence update)  
 T 28-FEB-2003 (Rel. 41, Last annotation update)  
 E Trophinin-associated protein (Tastin) (Trophinin-assisting protein).  
 N TROAP.  
 S Homo sapiens (Human).  
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 X NCBI\_TaxID=9606;  
 X [1]  
 P SEQUENCE FROM N.A.  
 P MEDLINE=95278733; PubMed=7758945;  
 A Fukuda M.N., Sato T., Nakayama J., Klier G., Mikami M., Aoki D.,  
 A Nozawa S.;  
 T "Trophinin and tastin, a novel cell adhesion molecule complex with  
 potential involvement in embryo implantation.";  
 L Genes Dev. 9:1199-1210(1995).  
 C -!- FUNCTION: COULD BE INVOLVED WITH BYSTIN AND TROPHININ IN A CELL  
 C ADHESION MOLECULE COMPLEX THAT MEDIATES AN INITIAL ATTACHMENT OF  
 C THE BLASTOCYST TO UTERINE EPITHELIAL CELLS AT THE TIME OF THE  
 C EMBRYO IMPLANTATION.  
 C -!- SUBUNIT: DIRECTLY BINDS BYSTIN, AND INDIRECTLY TROPHININ.  
 C -!- TISSUE SPECIFICITY: CYTOPLASMIC.  
 C -!- TISSUE SPECIFICITY: STRONG EXPRESSION AT IMPLANTATION SITES. WAS  
 C EXCLUSIVELY LOCALIZED TO THE APICAL SIDE OF THE  
 C SYNCYTIOEPHOBLAST. ALSO FOUND IN MACROPHAGES.  
 C -----  
 C THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C -----  
 C ENBL; U04810; AAA79333.1; --  
 C FIR; I38487; I38487.  
 C R Genew; HGNC:12327; TROAP.  
 C MIM; 603672; --

DR GO; GO:0005737; C:cytoplasm; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 KW Cell adhesion; Repeat.  
 FT DOMAIN 504 687 CYS-RICH.  
 FT DOMAIN 516 647 4 X 33 AA APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 516 548 1.  
 FT REPEAT 549 581 2.  
 FT REPEAT 582 614 3.  
 FT REPEAT 615 647 4.  
 SQ SEQUENCE 778 AA; 83758 MW; 52BAB17165672AC0 CRC64;  
 Query Match 22.2%; Score 131; DB 1; Length 778;  
 Best Local Similarity 22.4%; Pred. No. 0.1;  
 Matches 107; Conservative 44; Mismatches 167; Indels 160; Gaps 26;  
 Y 2 TRAPRCVAVSLRSRYREVWPLATFVRRLGPEGRLVQDPGPKYRILVA-----QC 54  
 Db 358 TPESTPRVQ-----AOWLRGVSPQS-----CSEDPALFWEQVAVRLFDQESC 401  
 Y 55 LVCMHWSQPP---PADLSFHOVSSLKELVAR---VVORLCERNERNVLAFGFELLNEAR 108  
 Db 402 IRSLEGCKPVPATPSGPHSNRTPSLOEVKIQBIGILOQLLQEQVEGLVGQCVPLN--- 458  
 Y 109 GGP--PMAFTSSVRSYLPNTVETLRVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSC 166  
 Db 459 GGSLLDMVELQLLLEISRTLNATEHNSGT-----SHLPGLLKHSGL-----PKPC 504  
 Y 167 AYQVCGSPLYCICATTDIW-----SYSASVRYRTPVGRNFTN-L 205  
 Db 505 LPEEGEP--QCPAPBPGPPEAFCSRSEPEIPSPSQEQLEVPPEYPPAEP--RPLESCC 560  
 Y 206 RFLQIKSSRQE-----APKPL-----ALPSRGTKEHLSLTSTSVPSAKK 246  
 Db 561 RSEPEIPSSSRQEQLEVPPECPAEPPLSEYCYRIPEIPSSRQEQ-----EVPE--- 612  
 Y 247 ARCPVPRVEGHRQVLPTPSGKSWVPSPARSPVPTAEKDLSSKGKVDLSLSGV--- 304  
 Db 613 ----PCPAPFPGP---LQPTQCGSGPPGFC-----PRVELGASEPTLEHRSLESSLP 660  
 Y 305 CCKH-KPSSTSL-----SPPRQNAQLPFPFETRFHFLYSRGDQGERLNPSFLSN 354  
 Db 661 CCSQWAPATSLFSSQHPLCASPPICSLQLRP-----PAGQAG-----LSN 703  
 Y 355 LQPNLTGARLV-----EIPLGRSPTSGP--LCRTHRLSRRYQ 393  
 Db 704 LAPRTLALRESLSCLTAIRHCFHEARLDDECAFYTSRASPSGPTRYCTNPVATLLEWQ 761

RESULT 9  
 DYHC PARTE STANDARD; PRT; 4540 AA.  
 AC Q27171;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Dynein heavy chain, cytosolic (DYHC).  
 GN DHC-8.  
 OS Paramaecium tetraurelia.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
 OC Paramaecium.  
 OX NCBI\_TaxID=5888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Stock 51;  
 RX MEDLINE=96157890; PubMed=8589455;  
 RA Kandl K.A., Forney J.D., Asai D.J.;  
 RT "The dynein genes of Paramaecium tetraurelia: the structure and  
 RT expression of the ciliary beta and cytoplasmic heavy chains.";  
 RL Mol. Biol. Cell 6:1549-1562(1995).  
 CC -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular  
 CC retrograde motility of vesicles and organelles along microtubules.  
 CC Dynein has ATPase activity; the force-producing power stroke is  
 CC thought to occur on release of ADP.

```

497 QY 497 ELM--WKWKVEDCHW-----LRSPGKORVPAERHLRERILATP--LF-----WLMDT 541
Db 514 RALKQYIAIESVEATITNLRDSLGS-----ASSAKEMFRILAKFNKLFSPRIRKGAIOE 569
QY 542 YVQQLASFFVITES-----TFQR--NRLFFYRK-----SVMSKLOSIGVROHLE 584
Db 570 YQSOLLKTVHKDIQSLQNKPEYIQKSONSRLASARDIPTLTSGFVINSQQLQIRLQKTMQ 629
QY 585 RV-----RLRELQ-----BEVRHQDTWAMPICRLRF 613
Db 630 KYEQILQPWAEDTDGKKCKCKEMGETFERILDSGPALEDWKQEIHNH-----KA 678
QY 614 IPKENGRLRPIVMSYSMGTRALGRRRQAOCHFTQRLKTLFSLNLYERTKHPHLMGSSVLGM 673
Db 679 VSQNEKLFVVTRRRGLIRVNYEKKSLQLF-KEVRNLSNM-----KTKVPYSISHIA-- 730
QY 674 NDIYRTWFAVFLVRALDQTPRMFYFKVADVTGAYDAIQGKLVEVVANMKRHSSESTYCIIR 733
Db 731 NDAKASY-PFAL---SLQESLHTYI---QITSQLNA-KSAKIV----- 765
QY 734 QVAVVRDSDSQGVHKSFRQVTLTSLDLPYMQOFLKHL----- 771
Db 766 --AALKREVLQIQGQENYLWHTKTLQYFVKFKTDKYFELEQAVNGLNRIQIESLCE 823
QY 772 -----QSDASALRNSVITEQISWNSSSLDFFLHFLRHSVVKIGDRCYTCOGCIP 825
Db 824 AMKTCPVDSLADKLKDIQEVDSLCFNNFSN-----LHIWTQIDK----- 864
QY 826 QGSSLTSLTSLSCFGDMENK-----FAEVORDGLL-----LRFVDFDLVTPLHDOA 873
Db 865 ---QIESILCDRTVQMKEWLNGFIYQKIQERGLVQTVVHELKLDQDIIVDPFPVEYA 921
QY 874 KTFILSTLVHGVPEYGCMMINLQKTVVN 899
Db 922 KYFWFOBFHKM--IGQICSLPRLVAN 945

RESULT 10
ID GEND_VIBU STANDARD; PRT; 873 AA.
AC Q8DEG3;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE [protein-PII] uridylyltransferase (EC 2.7.7.59) (PII uridylyl-
transferase) (Uridylyl removing enzyme) (UTase).
GN GEND OR VVI1857.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.-Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Modifies, by uridylylation or deuridylylation the PII
CC (GlnB) regulatory protein [by similarity].
CC -!- CATALYTIC ACTIVITY: UTP + [protein-PII] = diphosphate + uridylyl-
CC [protein-PII].
CC -!- SIMILARITY: Belongs to the glnB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isdb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC

```

ESIII.T 11





```
262 LINQLSDPDYINQVWIRDSNCNVEAFMNIKLSNDNIGELAEVRDKAAEL----- 314
538 LMDTVVQVLLRESFY-----ITESTFOKRLFFYRKSVWVKLOIGVROHLERVRRLAELSO 593
315 -----OYLSRLDTAGDINTIKQINSLLFVKVKCDRSIORLQSGKEINTVK----- 361
594 BEVRHHDOTWLMPICLRLRFPKNGRLRPIVNMYSM-----GTRA-----LGRK 639
362 -----LAANFGKLCITVPLDSILVDNALOFFMNDYMQOATGGQAHLLFFMWVGEYRV 411
640 QAOHFTQBLKLFSL-----NYERTKHPHLMGSSVLGMNDIYRTWRAFLVRALDQTPRM 696
412 TAQ-----OOLEVLRSQRDGHQNTQTKGLLRAAAGVIEQYLSEKA-----SPR- 457
697 YFVKADVTGAYDAIPQGLVEVAVNMIRHSESTYCI-----ROYAVVRDSOGOVHKSF 750
458 -----DYLVAKLADTLNHDPTPEIFDDIQRKVYELMURDE--RYPSPF 503
751 RRQ-----VTTLSLQPYMGQFUKHLQDSALRNSVVEIOSINNESS 796
504 QNALYVRMLAEMLDKDPSRGSDDGDSFGSPGTSINLSLDLSN 552
(1)
SEQUENCE FROM N.A.
TISSUE=Bone marrow;
MEDLINE=96281124; PubMed=8724849;
Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
"Prediction of the coding sequences of unidentified human genes. V.
The coding sequences of 40 new genes (K1AA0161-K1AA0200) deduced by
analysis of cDNA clones from human cell line KG-1.";
DNA Res. 3:17-24 (1996).
-- SIMILARITY: Contains 1 TTL domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; D79995; BRA11430.1; -
InterPro; IPR004344; Tub_tyr_lygase.
Pfam; PF03133; TTL; 1.
Hypothetical protein; Polymorphism.
DOMAIN 551 546
VARIANT 418 418 R -> H (IN dbSNP:2114664).
/FTID=VAR_013140.
SEQUENCE 1199 AA; 133376 MW; 5306DSFA37431081 CRC64;
```

```
Query Match
Best Local Similarity 18.3%; Pred. No. 3.3;
Matches 201; Conservative 140; Mismatches 320; Indels 437; Gaps 52;
Y 62 SQPPADLSFHVSLSKELVARVQRL-----CERNERNVL--AFGFELNARGGPPMAF 115
327 SQDPTKEIRF--TEAVRKLTAGFERKPRQCGQLQSSFLNPSFQWNLNRGRWKPPAV 384
```

```
QY 116 -----TSSVRSYLP-----NTVIETLRVSGAWMLLSRVGDDLLVYLAHAL 158
Db 385 NQOFPQEDAGSVRRVLPGASDTLGLDNTVFCTKRIS-----IHLASHAS 429
QY 159 YLLVPSQAYQVCGSPLYOICATTDIMPVSVASYPTRPVGRNFTNLFLOQIKSSSRQE 218
Db 430 GLMHNFACESEVI-----DSSAFGEKAGPPPF----- 456
QY 219 APXPLALPSRGTKRHLSTSTVPSAKKARCYPVPRVEGPHRQVLPTPSGKSWVPSPAR 278
Db 457 -POTLGIANVATR-----LSSIOIGQSEKER-----PEE 484
QY 279 SPRVPTAEKDLGSKGVSDLS-----LSG-SVCCKH-----KPSSTS 314
Db 485 ARELSDSDRDISS--ATDLQDQAEATEDIEELVDGLEDCSRDENEEBEGDSECSLS 541
QY 315 LLSPPFQNAQFLRPFLETBHFYLSRCDQOERLNPFLSNLOPNTLGGARRLVEIFLGSR 374
Db 542 AVSPSSVAMISKCM-----ILTKPLSNHEKVVRRPALLYSLPFNPV-----PTYFQT- 591
QY 375 PRATSGPLCRTHRLSRRYQWRPLFQ-QLLVNHAECQYVRLLRSHCRFTANQOVTDALNT 433
Db 592 -----RDERVEKLPEQKLLRWKMSVTVPNIVKOTIGRSHFKISKRND- 636
QY 434 SPPLMDLLRLHSSPMQVYGFRLACLCKVVSASLWGTNRHRRRPFKNLKKFISLKYGL 493
Db 637 -----W-----LGWCG-HHMKSPSFRSIREHOKLNF- 662
QY 494 SLQELMWKMKVEDCHWLRSPPG-----KDRVPAAEHLRERILATFLFLMDTVV- 544
Db 663 -----PGSFQIGRKORLWRNLNLSRMQSRFGKKEFPFPQSIFILLPD 702
QY 545 -QLLRFFVITSTFQKNRLFFYKSV-----WSKL-----QSIGVROHLERVRLE 590
Db 703 AKLLRWKAW---ESSRSQKWIVKPPASARGIGIQVTHKWSQLPKRPLLVQRYLHKPYLIS 759
QY 591 LSOEVRHHDOTWLMPICLRLRFPKNGRLRPIVNMYSMGTALGRKQAOHFTQRLKT 650
Db 760 GSKFDLRIVYVTSYDP---LRIYLFSDGLVRPASCYSPSKSLG----- 802
QY 651 LFSMLNYERTKHPHLMGSSVLGMNDIYRT-----WRAFLVRALDQTPRMVYFV 699
Db 803 -----NKFMLHTNYSVNKNAEYQANADEMACQGHKN-----ALKALWN---YLS 844
QY 700 KADVTGAYDAIPQGLVEVAVNMIRHSS--TYCIRQYAVVRDSOGOVHKSFRRQVTTL 757
Db 845 QKGVNS--DAIME-KIKOVVVVKTIISSPYVTSLLKMT--VRRPY--SCHELFQDFMLD 897
QY 758 SLLQPYM-----GQFLKHLQSDSALRNSV-VIEOSISMNESS 796
Db 898 ENLKPWYLEVNISPSLHSSPLDISIKGMIRDLNLAGFVLPAEADISSPSSCSSTT 957
QY 797 SLDFDFELHRLSVVKIGDR------YTCOGIGIQQSSSLTLLCSLCPGD 841
Db 958 SL-----PTSPGDKCRMAPEHVTAQKKKAYYLTKIPDQDFYASVLDVLTDPDD 1006
QY 842 -----MENKLPFAEVORDGLL-----LRFVDD---FLLVTHLDOA----- 873
Db 1007 VRLIVMEDEFSRRCQFERIFPSHISRLRPFEPFQPRYFNILTTQWEGKYHGNKLGVDL 1066
QY 874 -----KTEL-----STLVHGVPEYGCMINLQKTVVN----- 899
Db 1067 LRWCYKGFHMGVSDSAPVWSLPTSLLTISKDDVILNFAFSKSETSKLGKSSCEVSILL 1126
QY 900 -----FPVEPGLTGAAPY-----QLPAHCLFPWCGLLDTDTLVEF-CDY 939
Db 1127 SEDGTTPKSKTKQAGLSYPQKPPSSKSDSEDTSKEPS-----LSTQTLPIVKCS- 1175
QY 940 SGYAQTS-IXTSLTFQSV 956
Db 1176 ---GQTSRLSASSTFOSI 1190
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on on: February 4, 2004, 14:09:14 ; Search time 45 Seconds  
(without alignments)

6434.112 Million cell updates/sec

itle: US-09-042-460-2

fect score: 5901

quence: 1 MTRAPRCFAVRLSRGYRE.....TILKAAADPALSTDRQTIID 1122

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

tal number of hits satisfying chosen parameters: 830525

imum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

SPTREMBL 23: \*  
1: sp\_archea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriaph: \*  
17: sp\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	4751	80.5	1128	11 Q9QXZ4	Q9QXZ4 mesocricetu
2	3217.5	54.5	1069	4 QNG48	Qng46 homo sapien
3	3002	50.9	11	11 Q9JK99	Q9jk99 ratus norv
4	2729	46.2	514	11 Q9R266	Q9r266 mus musculu
5	2397.5	40.6	1191	13 Q9DE32	Q9de32 xenopus lae
6	2195	37.2	807	4 QNG6C3	Qng6c3 homo sapien
7	2125	36.0	795	4 QNG38	Qng38 homo sapien
8	1373	23.3	524	4 Q9UBR6	Q9ubr6 homo sapien
9	1086.5	18.4	523	4 Q94807	Q94807 homo sapien
10	780.5	13.2	1123	10 Q9SES99	Q9ses99 arabidopsis
11	777.5	13.2	1123	10 Q9SPU7	Q9spu7 arabidopsis
12	729	12.4	1259	10 Q9AU13	Q9au13 cryza sativ
13	686	11.6	1261	10 Q8LKW0	Q8lkw0 cryza sativ
14	550	9.3	104	11 Q9JLM1	Q9jlm1 mus musculu
15	450	7.6	1032	5 Q8KUB3	Q8kub3 euplores cr
16	443	7.5	823	5 Q8SQQ0	Q8sqq0 encephalito

17	420	7.1	1108	5 Q818Z6	Q818z6 sterkiella
18	418	7.1	895	5 Q8GRC5	Q8grc5 paramescium
19	412	7.0	1135	5 Q818Z7	Q818z7 sterkiella
20	384	6.5	867	3 Q9P8T3	Q9p8t3 candida alb
21	378.5	6.4	896	5 Q8M0Q8	Q8m0q8 paramescium
22	378	6.4	867	3 Q9P8T2	Q9p8t2 candida alb
23	320	5.4	85	4 Q9UN86	Q9uns6 homo sapien
24	277	4.7	73	4 Q9UNR4	Q9unr4 homo sapien
25	274	4.6	52	11 Q9R0B3	Q9r0b3 mus musculu
26	265	4.5	1474	5 Q962F9	Q962f9 cryptospori
27	231	3.9	960	5 Q9NCP5	Q9ncp5 giardia lam
28	147.5	2.5	2518	5 Q81EH2	Q81eh2 plasmodium
29	136.5	2.3	425	2 OS2231	OS2231 streptococc
30	134.5	2.3	418	16 Q8EQA2	Q8eqa2 oceanobacil
31	133.5	2.3	507	5 Q95QP3	Q95qp3 caenorhabdi
32	133.5	2.3	561	5 O45321	O45321 caenorhabdi
33	131.5	2.2	330	16 Q8DN10	Q8dn10 streptococc
34	130.5	2.2	436	2 Q9AE17	Q9ael7 streptococc
35	130.5	2.2	436	2 Q93DS9	Q93ds9 streptococc
36	130	2.2	425	16 Q8DZ79	Q8dz79 streptococc
37	130	2.2	778	4 Q8N5B2	Q8n5b2 homo sapien
38	129.5	2.2	419	16 Q8EMX5	Q8emx5 oceanobacil
39	129	2.2	502	5 Q22002	Q22002 caenorhabdi
40	128.5	2.2	369	16 Q8EP73	Q8ep73 oceanobacil
41	126.5	2.1	419	16 Q8EUL6	Q8eul6 oceanobacil
42	125	2.1	1137	11 Q8EIS9	Q8eis9 mus musculu
43	125	2.1	1529	10 Q81068	Q81068 arabidopsis
44	122	2.1	1192	11 Q8BI79	Q8bi79 mus musculu
45	121	2.1	783	3 Q9HE09	Q9he09 schizosacch

## ALIGNMENTS

RESULT 1

Q9QXZ4 PRELIMINARY; PRT; 1128 AA.

AC Q9QXZ4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Telomerase catalytic subunit.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21240330; PubMed=11342218;  
RA Guo W., Okamoto M., Lee Y.M., Baluda M.A., Park N.H.;  
RT "Enhanced activity of cloned hamster TERT gene promoter in transformed cells";  
RL Biochim. Biophys. Acta 1517:398-409(2001).  
DR EMBL; AF149012; AAF17334.1; -;  
DR InterPro; IPR000477; RVTse.  
DR InterPro; IPR003545; Telomerase\_RT.  
DR Pfam; PF00078; rvt. 1.  
DR PRINTS; PR01365; TSEMERASERT.  
KW RNA-directed DNA polymerase; Transferase.  
SQ SEQUENCE 1128 AA; 128394 MW; 1D4F51249012174E CRC64;

Query Match 80.5%; Score 4751; DB 11; Length 1128;  
Best Local Similarity 80.1%; Pred. No. 0;  
Matches 904; Conservative 90; Mismatches 127; Indels 8; Gaps 4;

QY 1 MTRAPRCFAVRLSRGYREVWPLATFVRLGEGRLVQPGDPKIYRTLVQAQCLVCMWH 60

DB 1 MPRAPRCFAVRLSRGYREVWPLATFVRLGEGRLVQPGDPKIYRTLVQAQCLVCMWH 60

QY 61 GSQPPADLSFHQVSSLKELVARVQVQLCERNERNVLAFGFELLNEARGCPPTTSYR 120

DB 61 DSQPPADLSFHQVSSLKELVARVQVQLCERNERNVLAFGFELLNEARGCPPTTSYR 120

```
2Y 121 SYLNTVITELRVSGAMLLLSRVGDDLLVYLAAHCAVLLVPPSCAYQVCGSPLYQICA 180
Db 121 SYLNSVTESLRVSGAMLLLNVRGDDLLVYLARCAVLLVPPSCAYQVCGSPLYQICA 180
2Y 181 TTDIWPSVASVYPTRVGVGNFTNLFLQIKSSSQEAPKPLAPSRGTRKHLSTGTS 240
Db 181 TAETWPSVSRVYPTRVGVGNFTHLGSTRVNSSHQEAKEKPPPLPSREAKESLSITNRS 240
2Y 241 VPSAKKARCYPVRVBEGRHQVLPPTPSGKSWVPSPARSPEVP---TAEKOLSSKGVKYSVD 297
Db 241 VPPSKKARCCLAPLEKRGPRVQVPTPSDKTWVNPAPKSHAVPIRSTTKEDLSSGVKAPG 300
2Y 298 LSLSGSVCCCKHKSSTSLSPQPNQAFQRPFTTRHFLYSSGDGERLNPSFLLSNLQOP 357
Db 301 LSRSGSVCYKHKSSSTSLQPLCNQAFQRPYETETKRFYLSREGGERLNPSFLLNNLQOP 360
2Y 358 NLTGARELVEIIFLGSRRPTSGPLCTHRLSRRYWQMRPLFOQLLVNHAEOYVRLLRSH 417
Db 361 SLTGARELVEILFLGMRPTSGPLCGRRRLSKRYWQMRPLFOQLLVNHAERCPYVRLLRSH 420
2Y 418 CRPTANQOYTDALN-TSPHMLDLRLHSSPMQVYGFELRACLCVKVSGASLWGTREHNR 476
Db 421 CRFTAAHQVAGALNTTSPRLNMLRLHSSPMQVYGFELQACVGVKLVPPGLMGSRRNQR 480
2Y 477 FFKNLKXFIISLGKYKLSLQELMWMKVEDCHWLRSPPGKDRVPAAEHRLRERILATLFL 536
Db 481 FFKNVKRFISLGKYKLSLQELTMKMKVQDCWLRSSPGNCVCPAAEHRTREILAVLFL 540
2Y 537 WLMDTYVQLLRFFYITESTFQKNLRFYRKSVWSKLOSIGVROHLRVLRLSQEVEV 596
Db 541 WLMDAYVVELLSFFYVYETFTFQKNLRFYRKSMWRRLQSIGVRHHLRVLRLQELSQEV 600
2Y 597 RHQDWTWLMPIICRLRFIPKNGRLPIVNNVSMGTRALGRKQAOHFTQRLKTLFSLMLN 656
Db 601 RQOEAWPMPICELRFIPKPSGLRPIVNNVSY-MGTRAFCKGKQAOHFTQCLKTLFSLVLN 659
2Y 657 YERTKPHLMGSSVLGNDIYRTWRAFLVRVRLALDQTPMYFVKADVTGAYDAIPQGLVL 716
Db 660 YELTKHTNLLGASVLGNDIYRTWRTFVLRVRLTDPAPMYFVKADVTGAYDAIPQDKVL 719
2Y 717 EVVANMIRHSESTYCIROYAVVRDVGOGVHKSFRQVTTLSLOPYMGOFKHLQDSDA 776
Db 720 EVIANMIRHFDNSYCIHQYAVVRDQGOQIHKSFRQVSTLSLOPHMGOFKHLQDSDT 779
2Y 777 SALNSVYVIBQSIISMNESSSLDFLHLFRLHSVKIGDRCTYQCQGIPOGSSSLTLCS 836
Db 780 SALNSVYVIBQSLSLNEASSSLDFFLRFVNSVVKIGRCVVCQGIPOGSSSLTLCS 839
2Y 837 LCFDGMENKLFABVQDGLLLRVDFDPLVTHLDOAKTFLSTLVHGVPEYGCMLNLOKT 896
Db 840 LCFDGMENKLFABVQDGLLLRFVDFLTVPHLVQAEAFRALVRGIPYEGCMLNLOKT 899
2Y 897 VVMPFVPGTGGAAPYQLPAHCLFPCWGLLLDTQTLFVFCDSYGAQTSIKTSLTFQSV 956
Db 900 VVMPFVADGLDGTAPHLPAHCLFPCWGLLLDTQTLFVLCYTGARTSIKASLTFTQT 959
2Y 957 FKAGKTWRNKLVLRLKCHGLFLDLQVNSLQTVNCINIKYIFLLQAYRFHACVQLPPDQ 1016
Db 960 FKAGRNMRKLLAVLRLKCHSLFLDLQVNSLQTVNCINIKYIFLLQAYRFHACVQLPPDQ 1019
2Y 1017 RVRNKLTFFLGITSSOASCCYATLKVNPNQWTLK---ASGSPFPPAAJWLCYQAFLLKLA 1073
Db 1020 HVRKNPAPFLLISINATSCYSILKVNAGMTLKAGAGSGSPFPPAAJWLCYQAFLLKLA 1079
2Y 1074 AHSVYIKCLGLPLTAQKLCRLKPLBATMTILKAAADPALSTDFQTILD 1122
Db 1080 GHSVYIKCLGLPLTAQKLCRLKPLBATMTILTAADPALSTDFQTILD 1128
```

RESULT 2

JEN646

CD Q8NG46

PRELIMINARY; PRF; 1069 AA.



```
181 TTIDWPSVSYRTPYGRNFTNLRFLOQIKSSRQEAQPKPLALPSRGTKRHLSTSTS 240
b 181 TTIDWPSVSYRTPYGRNFTNLRFLOQIKSSRQEAQPKPLALPSRGTKRHLSTSTS 240
y 241 VPSAKARCYPRVREBGPQVLPPTSGKSWPSPARSEVPTAEKDLSSCKGVSDLSL 300
b 241 VPSAKARCYPRVREBGPQVLPPTSGKSWPSPARSEVPTAEKDLSSCKGVSDLSL 300
y 301 SGSVCKKPKSPSTLSLSPRONAFOLRPFETRHLXSRGQGERLNPSPFLSNLOPNLT 360
b 301 SGSVCKKPKSPSTLSLSPRONAFOLRPFETRHLXSRGQGERLNPSPFLSNLOPNLT 360
y 361 GARRLVEIFLGSRRPTSGPLCRTHLSRRYQWRPILFOQLLVNHAECOVRLRSHCRF 420
b 361 GARRLVEIFLGSRRPTSGPLCRTHLSRRYQWRPILFOQLLVNHAECOVRLRSHCRF 420
y 421 RTANQOVTDALNTSPPHLMDLRLHSSPWQVYGFRLACLCRWASASLWGRHNERFFKN 480
b 421 RTANQOVTDALNTSPPHLMDLRLHSSPWQVYGFRLACLCRWASASLWGRHNERFFKN 480
y 481 LKXFTSLGKGLSLQELMWMKVEDCHWLRSPP 514
b 481 LKXFTSLGKGLSLQELMWMKVEDCHWLRSPP 514
RESULT 5
19DE32 PRELIMINARY; PRT; 1191 AA.
C QDE32;
D QDE32;
E 01-MAR-2001 (T-EMBLrel. 16, Created)
F 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
G 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
H Telomerase reverse transcriptase.
I TERT.
J Xenopus laevis (African clawed frog).
K Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
L Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
M Xenopodinae; Xenopus.
N NCBI_TaxID=8355;
O SEQUENCE FROM N.A.
P Kuramoto M., Ishikawa F.;
Q "Telomerase reverse transcriptase of Xenopus laevis.";
R Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
S EMBL; AF212299; AAC43537.1; --
T InterPro; IPR000477; RVTse.
U InterPro; IPR003545; Telomerase_RT.
V Pfam; PF00078; rvt; 2.
W PRINTS; PRO1365; TELOMERASERT.
X RNA-directed DNA polymerase; Transferase.
Y SEQUENCE 1191 AA; 138016 MW; 98D9D776869A57D6 CRC64;
Query Match 40.6%; Score 2397.5; DB 13; Length 1191;
Best Local Similarity 43.4%; Pred. No. 1.2e-193;
Matches 526; Conservative 212; Mismatches 341; Indels 133; Gaps 31;
2y 12 SLRSRYEYVPLATVRRLG-PEGRR--LVQGGDPKTYRTLVAQCLVCMWGSQPPAD 68
b 12 SLRLQYGVQLGIVEVTDLTQVGGIKVPVLGGDSKFRSFVABLVCIPRGTKPLSP 71
2y 69 LSPHOVSLXELVARVORLCERNENRNLAFGELLNEARGGPPNAFTSSVRSYLPNTVI 128
b 72 VSPQLQSTQREVVARVIOIRCKRNKNVAFGLVDE-KNSLRLTFNCFNPTT 130
2y 129 ETLRVSGAMLLSYGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICATTDIMPV 188
b 131 TTISTSILWETLLTRVGDDVVMVWLEQCSIFVFPVPCCYITGQPIYTL-PSDDVFLFQ 189
2y 189 SASVRPT----RPVGRNFTNLR--FLQOKSSS-----RQEAQPKPLALPSR----- 228
b 190 SQSFQSNVLLRYIKRNVFHLRKKYLPKXHSMTSRLTWRRNKSPSGLLIRSKTSMVTT 249
```

```
229 --GTXRHLSTSTSVSAKAR-----CYPVPRVEGPHR 261
250 EIHKRLKCSKDCVTPDKRRDNLKDDTVDFDLPMSRSVSYLSNTPKTNV----- 303
262 QV--LPTPSGKSWPSPARSEVPTAEK-----DLSSCKGVSDLSLSSGSCCKH 308
304 QVTGLITSGYKTKTQCQKPVSCQKKTAFYSVAGDCNLSLKONVKNKIITNASV---- 359
309 KPSSTLSLSPRONAFOLRPFETRHLXSRGQGERLNPSPFLSNLOPNLTGARRLVEI 368
360 -PTAQSRLS-----FS-NIFDIFGRTLYSISYKGFSEFILNSLSDTPSGSKLVET 411
369 IFLGS-----RPTSGPLCRTHLSRRYQWRPILFOQLLVNHAECOVRLRSHCRF 420
412 IFLNFAEQNFDQFED-ENCR-YKLPKRYWKKHFOELQNHKKFYLYLVNKHCFV 469
421 RTA-----NQOVTDALN-----TSPPHLMDLRLHSSPWQVYGFRLACLCRWYSA 465
470 RSSMACSEKRSLOKNRIENDGKQLKHFTTKANLLSLKQHSSTWQVYMEVRECLNNVVD 529
466 SLWGRHNERFFKNLKKFISL-GKYGKLSLQELMWMKVEDCHWLRSPPGKDRVPAAEH 524
530 IMWGSSENKCRFFRNKSFLLPFGKFKLSLSELMMSRVEDCSWIRLQKSHFVPASEH 589
525 RLRRILATFLWLMDTYVQVLLRSPFYITESTFOQRLFFYRKSVMSKLSQSIGVRQHLE 584
590 LLREKILAKFVWLMDTYV:QLLKSFPYVTTWFOQHRLLFYRKSVMSKLSQSIGVRQHLE 649
585 RVLRLBELSQEVRHODTWLAMP-ICELREIPKNGURPIVNMVSYMGTRAL--GRRKQA 641
650 KYKLSRSSDELENMQC-WKNVPLVSLRPIKTNGLRPLSKISSTLSQSQSKENGEKKI 708
642 QHFORLKTLPMLNRYERTKPHLMGSSVLGMNDIYRTWRAFVL-----RVRALDQTPRM 696
709 HFFSSQIENLPSVLNENENCSLIGSSVFGMDIYKWKKFVLDPEKQVEKL-----QF 764
697 YFVKADVTGADAIPOGKLVFVWVANMIR-HSESTYCIQYAVVRRDSQGVHKSFRQVT 755
765 YFVKTDVKGAYDTIPHSKLDVSKVINPNANEVYCIIRYATVSDPTGRIIKSFGRHVS 824
756 TISDLQVYMGQFLKHLODSASALRNSVWIEQSIWNESSSLDFPFLHFLRHSVVKIGD 815
825 ELADVLPNKKQFVSNQEK--NLLRNLTVQENLLNESSVKLLAVFQIIRSHILRIKD 882
816 RYTCQCGIPIQSGSLSTLLCSLCPGDMENKLFABVRDGLLLRFDVDDFLLVTPHLDQAKT 875
883 RYTMQCGIPIQSGSLSTLLCSLCPGDMENAMLGIOKNGVLMRLIDDFLLVTPHLDQAKT 942
876 FLSTLVHGVPEYGGMINLOKTVNFPVE--PGTLGGAAPYQLPAHCLFPWCGLLDQTL 933
943 FLRTLAEGIPQYGCISFPQKTVNFPVDDIPEC---SEVEQLFSLCLFRWCGLLDQTL 999
934 EVFCDYSYGAQTSIKTSLTFSQVFAGKTWRNKLVSRLKCHGLFLDLQVNSLQTVGIN 993
1000 DIVYDYSYACTSIRSSMTFCHSSAAGKWKQKLIYRLKCHSLFLDLKVNLSRTVCIN 1059
994 YKIFLLQAYFHACVQLPDPQVRNLTFFLIGTSSQASCCYAILKVKNPQMT---LK 1050
1060 TYKIFLLQAYFHACVQLPDPQVRNLTFFLIGTSSQASCCYAILKVKNPQMT---LK 1119
1051 ASGSPFPPEAAHWLCYQAPFLKLAHVSIVYKCLGLPLTAQKLCRKLPEATMTILKAAAD 1110
1120 VSCQNFENAVQWLSYQAFLTKLHNHVKLYKCLGLPLQCNKQMLSRRLSQDTILLKSVTD 1179
1111 PALSTDFOTILD 1122
1180 SSLHKDFSCIMD 1191
RESULT 6
Q8N6C3 PRELIMINARY; PRT; 807 AA.
ID Q8N6C3
AC Q8N6C3;
```



```

01-OCT-2002 (TREMBlrel. 22, Created)
01-OCT-2002 (TREMBlrel. 22, Last sequence update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
Beta and gamma deletion isoform of telomerase reverse transcriptase.
HTEPT.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Stomach cancer;
Hisatomi H., Nagao K., Kanamaru T., Sumida H., Hirata H., Yamamoto M.,
Kazumasa H.;
"Both beta and gamma deletion isoform of human telomerase reverse transcriptase.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AB086379; BAC11014.1; -.
InterPro; IPR003545; Telomerase_RT.
PRINTS; PR01365; TELOMERASERT.
SEQUENCE 807 AA; 90225 MW; 199664460CE6D763 CRC64;

Query Match 37.2%; Score 2195; DB 4; Length 807;
Best Local Similarity 58.3%; Pred. No. 9.7e-177;
Matches 462; Conservative 88; Mismatches 191; Indels 52; Gaps 12;

1 MTEAPPCAVRSLLRSRYREVWPLATFVRLGPEGRLVQPDGPKIVRTLVAOCLVCMHW 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MPRAPPCAVRSLLRSRYREVWPLATFVRLGPEGRLVQPDGPKIVRTLVAOCLVCMHW 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

61 GSOPPPADLSFHQVSSLSKELVARVYQRLCERNVLAFFGELLNARGGPPMAFTSSVR 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 DAREPPAAFSFQVSLKELVARVYQRLCERNVLAFFGELLNARGGPPMAFTSSVR 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

121 SYLPNTVITLTVSGAMWLLSRVGGDLLVYLLAHCALYLLVPSCAYQVCGSPLYQICA 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 SYLPNTVITLTVSGAMWLLSRVGGDLLVYLLAHCALYLLVPSCAYQVCGSPLYQICA 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

181 TTDIWPSVASYRPTPVGRNFTNLRFLOQIKSSROEAPKPLALPSRGTKRHLSTSTS 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAGARRRGGASRS 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

241 VPSAKKARCYFVPRVEEGP-----HRQVLTPSGKSW-VPSPARSEPVTAEKDLSK 292
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 LPLPKRPRGAAPERTPTVQGGWAHPGRTGSDRGFCVVSAPAR-----PABEATSLE 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

293 GKVSDDLSS-GSVCCCHKPSSTLSPPRQNAFQLRP-FIETRHFLYSRGDOERLNPSP 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 GALSGTRHSHPSVGRHAGPPSTRPRPMDTPCPPVYAEKTHFLYSSGD-KEQLRPSF 345
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

351 LLSNLOPNTGARRLVEIIFLGSRRPTSGPLCRTHRLSRRYQWQRPFLQOLLVNHACQY 410
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 LLSSLRPSLTGARRLVEIIFLGSRRPTSGPLCRTHRLSRRYQWQRPFLQOLLVNHACQY 410
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

411 VLLRSHCHFRPTANQOVTDAL-----NTSPHLLMDLLRLHSSPMQVY 452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
406 GVLLKTHCHPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRLVLQRLHSSPMQVY 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

453 GFVRACLCKVVSASLWGTNRHRRFPNKKFISLGKYGKLSLOELMMKMWEDCHWLRS 512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 GFVRACLRLVPPCLWGSRRHNERFLNKKFISLGKHAKLSLQELTWQSVRCAMLRR 522
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

513 SPQKDRVPAEHLRLREILATFLWMDTVYVQLLRSFFYITESTFQKRLFFYKRVWS 572
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 SPQGVGCVPAEHLRLREILAKELHLMNSVYVVELLRSFFYVTEFTFQKRLFFYKRVWS 582
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

573 KLOSIGVROHLERVLRELSQEBEHQDTWAMPICRLRTPKENGHRPIVNNYSNGT 632
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
583 KLOSIGVROHLERVLRELSQEBEHQDTWAMPICRLRTPKENGHRPIVNNYSNGT 642
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

633 RALGRKQQAQHTQRLKTLFSLMNYERTKHPLMGSSVLGMNDIYRTWRAFLVRALDQ 692
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

b 463 GFVRACTRLVPPGLWGRHNRERFLNRTKFKFISLGHAKLSQELTWGNSVRDCAWLRR 522
y 513 SPGKDRVPAARHLRERLATLFLWMDTYVQLRSFFYITESTFQKRLFFYRKSVMS 572
c 523 SPGVCVPAARHLRERLATLFLWMDTYVQLRSFFYITESTFQKRLFFYRKSVMS 582
y 573 KLOSGVQHLRERVLRLSQRVHHQDTWAMPICRLRFIPKPNGLRPIVNMYSMGT 632
c 583 KLOSGIGIRHLRERVLRLSQRVHHQDTWAMPICRLRFIPKPNGLRPIVNMYSMGT 642
y 633 RALGRKKAQOFTORLTKLFLMNLVYRTKPHLMGSSVLQNMNDIYRTWRAFLVRVRLDQ 692
b 643 RTFREKEAERLTSRVKALFSLVNYERARRPGLLGASVGLGDDIHRARTFVLVRAQDP 702
y 693 TPRMVFVADVTGAYDAIPQOKLVEVNVANMRHSESTYCIQVAVVRDSDQGVHKSFR 752
b 703 PPLEYFVK-----DRLTEVIAIHK-PONTYCVRRYAVVQKAHGHVRKAFKS 749
y 753 QVTTLSLDLPYMG 765
b 750 HV-----LRPVP 757

RESULT 8
9UBR6 PRELIMINARY; PRT; 524 AA.
D Q9UBR6
T Q9UBR6;
T 01-MAY-2000 (Tremblrel. 13, Created)
T 01-MAY-2000 (Tremblrel. 13, Last sequence update)
T 01-MAY-2000 (Tremblrel. 13, Last annotation update)
E Telomerase reverse transcriptase (Fragment).
TERT.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
N [1]
P SEQUENCE FROM N.A.
X MEDLINE=99144726; PubMed=10022128;
A Greenberg R.A., O'Hagan R.C., Deng H., Xiao Q., Hann S.R., Adams R.R.,
A Lichtsteiner S., Chin L., Morin G.B., DePinho R.A.;
T "Telomerase reverse transcriptase gene is a direct target of c-Myc but
T is not functionally equivalent in cellular transformation.";
L Oncogene 18:1219-1226 (1999).
N [2]
P SEQUENCE FROM N.A.
X MEDLINE=99140777; PubMed=9988278;
A Wu K.J., Grandori C., Anacker M., Simon-Vermot N., Polack A.,
A Lingner J., Dalla-Favera R.;
T "Direct activation of TERT transcription by c-MYC.";
L Nat. Genet. 21:220-224 (1999).
L R EMBL; AF121948; RAD24464.1; -.
W EMBL; AF114847; RAD17210.1; -.
R RNA-directed DNA polymerase.
T NON TER 524
T SEQUENCE 524 AA; 57932 MW; 5F47DEFD01832B1B CRC64;

Query Match 23.3%; Score 1373; DB 4; Length 524;
Best Local Similarity 56.1%; Pred. No. 2.1e-107;
Matches 304; Conservative 50; Mismatches 142; Indels 46; Gaps 10;

y 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPRGRILVOPGDPKIVRTLVAOCLVCMHW 60
b 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPRGRILVOPGDPKIVRTLVAOCLVCMHW 60
y 61 GSQPPADLSFHQVSSLKELVARVQVRLCERNERNVLAFCGELLNEARGGPPMAFTSSVR 120
b 61 DARPPAAPSPQVCSLCELVARVQLCERGAKNVLAFCGALLDARGGPPPEAFTTSVR 120
y 121 SYLNTVITELRVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYQVCGSPLYQICA 180
b 121 SYLNTVITELRVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYQVCGSPLYQICA 180
y 181 TTDIWPVSASVYRTPRVGNFTNLRFLOQIKSSSRQAPKPLALPSRGTKHLSTSTS 240
b 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGPAFGARRRGSGASRS 231

```

```

QY 181 TTDIWPVSASVYRTPRVGNFTNLRFLOQIKSSSRQAPKPLALPSRGTKHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGPAFGARRRGSGASRS 231
QY 241 VPSAKKARCVPVPRVEEGP-----HRQVLTPSGKSW-VSPSPARSPEVPTAEKDLSSK 292
DB 232 LP.LPKRPRCAAPERTPTVGQSGMAHPGRTGRPSDRGFCVVSPAR-----PAEATSLE 286
QY 293 GKVSDDLSSLS-GSVCCGHKPSSTLSLSPRONAFQLRP-FIETRHFLYSRGQOERLNSF 350
DB 287 GALSGRTHSPSVGRCHAGPPSTSPRPWDTPCPVPVYAEATKFLYSGD-KEQLRPSF 345
QY 351 LLSNLQNLTGARRLVEIIFLGSRPRTSGPLCRTRLRSRYWQMEPLFOOLLVNHAECQY 410
DB 346 LLSLRSLTGARLVETIFLGSRPWMPGTPRRLPLPORYWQMEPLFLELLGNHAQCPY 405
QY 411 VLLRSRCRFTANQVTDAL-----NTSPHMLDLLRLHSSPWQVY 452
DB 406 GVLLKTHCP.LRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRRLVQLLRQHSFPQVY 462
QY 453 GFLRACLCKVVSASLWGTNRNRRFPKLLKFLISLGYKLSLQELMWKQKVEDCHWLSR 512
DB 463 GFVRACTRLVPPGLWGRHNRERFLNRTKFKFISLGHAKLSQELTWGNSVRDCAWLRR 522
QY 513 SP 514
DB 523 SP 524

RESULT 9
O94807 PRELIMINARY; PRT; 523 AA.
AC O94807;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Telomerase transcriptase (Fragment).
GN HTERT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takakura M., Kyo S., Kanaya T., Hirano H., Takeda J., Yutsudo M.,
RA Inoue M.;
RT "Cloning and characterization of human telomerase catalytic subunit
RT (hTERT) gene promoter.";
RL Cancer Res. 0:0-0 (1999).
DR EMBL; AB016767; BAA74724.1; -.
FT NON TER 523
SQ SEQUENCE 523 AA; 56555 MW; 8FDE562DDECC93DA CRC64;

Query Match 18.4%; Score 1086.5; DB 4; Length 523;
Best Local Similarity 48.7%; Pred. No. 3.9e-83;
Matches 269; Conservative 45; Mismatches 171; Indels 67; Gaps 16;

QY 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPRGRILVOPGDPKIVRTLVAOCLVCMHW 60
DB 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPRGRILVOPGDPKIVRTLVAOCLVCMHW 60
QY 61 GSQPPADLSFHQVSSLKELVARVQVRLCERNERNVLAFCGELLNEARGGPPMAFTSSVR 120
DB 61 DARPPAAPSPQVCSLCELVARVQLCERGAKNVLAFCGALLDARGGPPPEAFTTSVR 120
QY 121 SYLNTVITELRVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYQVCGSPLYQICA 180
DB 121 SYLNTVITELRVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYQVCGSPLYQICA 180
QY 181 TTDIWPVSASVYRTPRVGNFTNLRFLOQIKSSSRQAPKPLALPSRGTKHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGPAFGARRRGSGASRS 231

```

```

241 VPSAKARCYVPVREBGP-----HQVLPSPGKSW-VPSPARSPVPTAEKOLSK 292
232 LFLKPRRGRGAPEPERFVQGSWAHFGRTGRGSDRGFCVVSAPAR-----PABEATSLE 286
293 GKVSLSLS-GSVCCKHKPSPSTLSLPPRONAFOLRP-FIETRHFLYSRGDGOERLNPSF 350
287 GALSOTRSHSPSVGQHAGPSPSTSRPRPWTCTCPPVYAEKHFLYSSGD-KEQLRPSP 345
351 LLSNLOPNLTGARRLVEIFLGSRRPTSG-----PLC--RTHLRSSRYWQMRPLFQOLLV 403
346 LLSLRPLTGTARRLVEIFLGSRRPTSG-----PLC--RTHLRSSRYWQMRPLFQOLLV 403
404 NHAECQVRLRLSHCRFTANQOVTDALNTSPHMLDLRLHSS---PWQVYGLRACLC 460
406 G---CSSRRAR--CELASPPQVPS-VPGRSFRALWRRPRRTOTPVAM-----CSC 451
461 KVSASLWGR-----HNERFFKNLKKFISLGYKLSLOELMWKM 502
452 SASAPLAGVRLRAGLAPAGAPRPLGLQAORRFLRNTKKFISLGHAKLSLOELTWKM 511
503 KVEDCHLWLRSSP 514
512 SVRDCAWLRSSP 523

SULT 10
SE99 PRELIMINARY; PRT; 1123 AA.
Q9SE99
Q9SE99:
01-MAY-2000 (TREMELrel. 13, Created)
01-MAY-2000 (TREMELrel. 13, Last sequence update)
01-MAR-2003 (TREMELrel. 23, Last annotation update)
Telomerase reverse transcriptase catalytic subunit.
TER.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
STRAIN=landsberg erecta;
MEDLINE=20079564; PubMed=10611295;
Fitzgerald M.S., Rika K., Gao F., Ren S., McKnight T.D., Shippen D.E.;
"Disruption of the telomerase catalytic subunit gene from Arabidopsis
inactivates telomerase and leads to a slow loss of telomeric DNA.";
Proc. Natl. Acad. Sci. U.S.A. 96:14813-14818(1999).
EMBL; AF172097; AAD54276.1; -.
InterPro; IPR000477; RVTse.
InterPro; IPR000215; Serpin.
InterPro; IPR003545; Telomerase_RT.
Pfam; PF00078; rvt; 1.
PRINTS; PR01365; TELOMERASERT.
PROSITE; PS00284; SERPIN; 1.
RNA-directed DNA polymerase; Transferase.
SEQUENCE 1123 AA; 130551 MW; 44248295817B7F6 CRC64;

Query Match 13.2%; Score 780.5; DB 10; Length 1123;
Best Local Similarity 24.1%; Pred. No. 1e-56;
Matches 290; Conservative 208; Mismatches 435; Indels 269; Gaps 42;

38 LVQPGDPKIXRTLVACQVCMHWSQPPADLSFHOVS--SLKELVARVVORL---CERN 92
60 LKRSDDPHYRKLHRCFVVLH--EQTPPL-LDFSPTSWWSQREIVERIEMQSGCDC- 115
93 ERNLAFQFELLNEARGPPMAFTSVRSYLPNTVIEITLVSGAMMLLSKRGDDLLVYL 152
116 -QNVICARYD-----KYDQSSPILELTSSSWEPFLKRVGHDVAVVL 156
153 LAHCALYLLVPPSCAYQVCGSPLVQICNTTDIWSVSASVSRPRPVGRNFTNLFLOQIK 212
157 LQOTSIFLPLGLKHKHQVSGGPP-----CIKHKRTLSVHEN----- 192

```

```

213 SSSRQEPKPLALPSRQTKGH-LSLTSTVSPSAKARCYVPVREBGPVPTAEKOLSK 271
193 KRKEDDNNQP-----PTKRWLSSAVDDCPKDDSATITPIVGEDVDQHKRKTTKRSRI 246
272 WPSPASPSPEVTAEDKL-----SSKGKVS-----DLSLGSVCCKHKPSPSTLSL 316
247 YLKERRKQRKVNPKVDCNAPCITPTNGVKVSTGNDENMLHIGINGSL-----TDV 298
317 SPPQNAQRLPFIETRHFLYSRGDGOERLNPSFLSNLOPNLTGARRLVEIF----- 370
299 K-----QAKQVRKNKVPFGLSETYSVIPPNHILKTLRPNCSDKLMMHIFGEVNV 351
371 -----LGSRPRTSGPLCRTHLSRRYQMRPLFQOLLVNHAEQVVELLSHC----- 418
352 STTFSHGKGNCP--SGSICUYSLLK---SLKNLIGTKSSHLK-----MLDKHCPVLL 402
419 -----RFTANQOVTDALNTSPPH-----LMDLLRLHSSPWQVYGLR 456
403 QEDALKSGTTSQSSRRQKADKL---PHGSSSSQCTGKPKCPSVEERKLYCTNDQVVSPI 458
457 ACLCKVYSASLWGRHNERFFKNLKKFISLGYKLSLOELMWKM 506
459 AICRYIIPESILGTTHQMRVLRKXIAWFSRRRNEKCTVNOFLHKVYKPSDFPFPAKELC 518
507 C---HMLRSSPGKDRVPAAEHRLRERILAT-FLFWLMDTYVYVQLLRSPFYITESTFO 561
519 CMVNGHELOSE-----SIRSTQCMCTKWSLFWLEIVKVLVHFNFYATESQGRL 569
562 RLFPYKSVMSKJQSIGVROHLRERVLRELSEBVEVHHQDTWLAMPICRLRIPKENGRL 621
570 NIYYRKRSWERLTSKEISKALDGYVLVDDAEASRRK-----LSKRFELPKANGVR 622
622 PIVNMSYMGTRALGRKQAOHRTKTLFSLMNYERTKHPHMGSSVLGMNDIYRTWR 681
623 NVLDFSSSR-----SQSLRDTHAVLKDIQLEKPDVLGSSVDFDHDYRNL 669
682 AFVLVRALD-QTPRMVYKADVTGAYDAIPQGLKEVVVANMRHSESTYCIQYAVV--- 738
670 PYLIHLRSQSGELPPLYFVADVAFKAFDSVDQGLLHVIOFLK---DEVILNCRCLVCC 726
739 -RRDS-----QGVHKSFRQVTTLSLDLOPYMGQFLKHLQDSASALRNSVITEQSI 790
727 GKRNWNWKNILVSSDRKNSNFSRFTSV---PY-----NALQ-SIVVDKGN 768
791 MNSSSLFDFFLHFLRHSVVKTCDCRYTCCQGIPOGSSLSLTCLSCFGDMENKL---- 846
769 HRVTKDLMVMWIGNMLKNMQLDKSPFYQIAGIPQGHRLSSILCCFYVYGHLETLIYFF 828
847 PAEVQORD-----GLLLRFVDDFLVTPHLDQAKTFTSLVHGVPEYGC 889
829 LEEASKDVSSKECSREELIPTSXYKLLRFIDYLFVSTSRDQASSFYHRLKHGFKDNC 888
890 MINLQKTVNFFVPEPGTGGAAAPYQPAHCLF-----PWCGLLDTQTLVEVFCDS 940
889 FMNETKCFINE-----DKBEHRCSSNMFMVGNGVFPVWRTGLLINSRTEFVQVDT 941
941 GYATSIKTSLTTSQSVFKAGKTWRNKLKLSVLRLKCHGLFLDQVNSLOTVCIYKIFLL 1000
942 RYLSGHISSTFSVWQNKVPVNLKCYFLVPKCHPILFDSNINSGEIVRLNIYQIFLL 1001
1001 QAYEFHACVQLPDDQVRKMLTF-FLGIISQASQCYAILKVKPMPGMLTKASGSPFP-- 1057
1002 AAMKFCYVTEVSFRWKLHPOTLKFITI-----SVYMFRLNRRVRINTSGSFRPVL 1056
1058 ----EAAHW-CYQAFLLKLAHSAVYIKCLLGPLRTAQKLCRKLPEATWTILKAADPAL 1113
1057 KLYKEEIVLGLDAYIQVLKKNRVRMILLIYLKSA--LSKHSLSQLSSELKLYATDRSN 1114
1114 ST 1115
1115 SS 1116

```

```

352 STTPSHKGNCP--SGSICLYHSLK---SLKNLIGTKTSSHLK-----MLLDKHPVLL 402
419 -----RFTANQVTDALNTSPH-----LMDLLRLHSSPWQVYGFRL 456
403 QEDALKSGTTSQSSRRQKADKL---PHGSSSSQTGPKCPSVEERKLYCTNDQVFWFIW 458
457 ACLCKVVSASLWGTNRHNRFPKLNKKFISLGKYGKLSLOELMMKMKVED----- 506
459 AICRYIVPESLIGTTHQVRLKNTAFVSRNRNEKCTVNGFLHKVPSPDFFPARKELC 518
507 C-----HMLRSSPKORVPAEHLREILAT-FLFWMMDTVVQLLSFFVITESTQKN 561
519 CMVNGHELQSE-----SIRSTQOMLCTKWISWLFLEIVVKLVHFNFYATSOQGR 569
562 RLFFYRKSVSKLOSIGVQHLERVLRLSELQEEVRRHODTWLAMPICRLRFIPKPNGLR 621
570 NIYYRKRWERLISKEISKALDGVVLDAAESSRK-----LSKFLPKANGVR 622
622 PIVNMSYSGTRALGRKQAOHFTQRLKTLFMSMLNYETKPHLMGSSVLGMNDIYRWR 681
623 MVLDFSSSR-----SQSLRDTHAVLKDIQKKEPDLVGLSSVPDHDFFYNLC 669
682 AFVLRVRLD-QTPMYFVKADVTGAYDAIPOGKLVEVYVANKIRHSESTYCIQYAVV-- 738
670 PYLHLRSQSGELPPLYFVADVFAKFDVQDGLLHVIOQSLK---DEYILNRCLVCC 726
739 -RRDS-----QGQVHKSFRQVTTLSDLQPMQOFLKHLQDSASALRNSVIEQIS 790
727 GKSNWVKILYSSDKNSFSFTSTV---PY-----NALQ-SIVVDKGEN 768
791 MNESSSLDFDFLHFLRHSVVKIGDRCTYQCOGIPQSSSLTLLCSLCFGDMENKL----- 846
769 HRYRKDLMVWIGNMLKNNMLQDKSFYVQIAGIPQGHRLSLSLCCFYGHLETLIYFF 828
847 FAZVQD-----GULLRFVDDELVTPLHLDQAKTFLSLTVHGVPEYGC 889
829 LEASKDVSKESCEEBELIIPTSYKLLRFIDYLFVSTRDQASSFYHRLKKGFKDYN 888
890 MINLQKTVNVFVPEOTLGGAAPOYQPAHCLF-----PWCGLLDTOTLEVCDS 940
889 FMNETKFCINFE-----DKEEHRCSNRMFVGONGVFPVFWTGLLINSRTFEVDYT 941
941 GYAQTSIKTSLTFQSVFKAGTMRNKLVLKCHGLFLDLQVNSLQTCVINYKIFLL 1000
942 RYLSGHISSTFVAMQKPVRLRQKCYFLVPKCHPILFDSNINSGEIVRLNIYQIFLL 1001
1001 QAVRFACVQLPFDQVRKNTLF-FLGISSQASCYAILKVNKPNGMTLKASGSPFP-- 1057
1002 AAMKFCYVYVSRFWKLPQTLFKFITI-----SVRYMFLINRRVRRINTGSSFRVL 1056
1058 -----EAAHWLCYQAFLLKLAHSAVYKLLGPRTAQKLLCEKLPBETMTILKAADPAL 1113
1057 KLYKEEVIWGLDAYIQVLKKNRYRMLLIYKGA--LSKHSLSQQLSSELRYATDRSN 1114
1114 ST 1115
1115 SS 1116

RESULT 12
Q9AUL3 PRELIMINARY; PRT; 1259 AA.
AC Q9AUL3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Telomerase reverse transcriptase.
GN TEXT.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;

38 LVOPGDPKIYRTLVACQVCMHWSQPPADLSPHQVS--SLKELVAVYORL---CERN 92
60 LLRSDDPHYKXLLHRCFVHLH--EQTPPL-LDPSPTSWWSQREIVERIIEWMQSCDC- 115
93 ENVTAFGELLNEARGGPPMAFTSVRSVRLPNTVETLAVSGAMMLLSRVGDDLLVYL 152
116 -QNVICARYD-----KYDQSPILELTSSTSSWFLKRVGHDVWVYL 156
153 LAHCALYLLVPSCAYQVCGSPLQICATTIMFSVSASVPRTPVGRNFTNLRFIQQIK 212
157 LQOTSIFLPLGKHQQVSGPPL-----CIKKRILSVHEN----- 192
213 SSSRQAEKPLALPSRGTKRH-LSLTSTSVESAKKARCYVPVPRVEGPHRQVLPTPSGKS 271
193 KKRKDDNVQF-----PTKRQWLSSAVDDCPKDSATITPIVGEDVDQHKREKTKRSRI 246
272 WTPSPARSPVPTAEKDL-----SSKGKVS-----DLSLSQSVCCCKHPSSTSL 316
247 YLKRKRKQKVNFKVDCNACFCITPSTNGKVSCTGDNMLHIGINGSL-----TDFV 298
317 SPPRQNAQFPIETRIHFLSRGDSGRNLNPSFLLSNLQNLGTGARLVEIIF----- 370
299 K-----QAKQVKNKFKGLSETSVIIPNHILKTLRPNCSDKLLNHHIFGEVNW 351
371 -----LGSRPRTSGPLCRTHLSRRYQWQRPFLPQQLLVNHAEQVILLRSHC----- 418

```

SEQUENCE FROM N.A.  
Oguchi K., Tamura K., Takahashi H.;  
"Molecular cloning and characterization of OstERT, a telomerase  
reverse transcriptase homolog in *Oryza sativa*."  
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
EMBL: AF288216; AAK35007.1; -.  
GenBank: Q9AU13; -.  
InterPro: IPR001209; Ribosomal\_S14.  
InterPro: IPR000477; RVTse.  
InterPro: IPR003545; Telomerase\_RT.  
Pfam: PF00078; rvt; 2.  
PRINTS: PR01365; TELOMERASERT.  
PROSITE: PS00527; RIBOSOMAL\_S14; 1.  
RNA-directed DNA polymerase; Transferase.  
SEQUENCE 1259 AA; 143710 MW; 158041789FD5CAD CRC64;  
Query Match 12.4%; Score 729; DB 10; Length 1259;  
Best Local Similarity 24.2%; Pred. No. 2.7e-52;  
Matches 297; Conservative 182; Mismatches 429; Indels 318; Gaps 47;  
4 APR-----CPAVRSLRSRYREVWPLATFVRRLGPEGRRLVQDPGPKYRTLV--AQCIV 56  
41 SPRHSCSPACGRVAGS-----CLACRR-----WEHLRDPPIAYRLITRAVCAI 89  
57 CMHWGSGPP-----PADLSFQVSSYLKELVARVQRLCERNERNVLAFGPELLNEARGGP 111  
90 AADLSAPPPTPTGNSGHSQARLVREMKKSIADQ-SHGTKNVLCNGLH-----EGGQ 143  
112 PMAFTSVRSLNPTVETLRVSGAMWLLSRVGGDLLVYLIAHCAVLLVPPSCAYQVC 171  
144 SIC-----ISDLVSSSSSILLHIGDMLLCYLLRCTSIPLVKQNDYFQVS 190  
172 GSPLYQI-----CATTDIWP-----VSASY 192  
191 GVPLNVLRNPIFASIVARKHQPTTKAKCHTCYILWKAENALSIChSDSSNGVNSSF 250  
193 RPRPV--GNFTNLRFLOQIKSSSRQEAPEKPLALPSRG-----TKRHLSTS 238  
251 SSTCKIVTQSCETCGSIRRAESKDPSEGCNCPKFPDGRSGECCNCTHNRKRKLYS 310  
239 TSVPSSAKACVYVPAVE-----GPHRQVLPFPSSOKSVVFPSPASPE 281  
311 WQRSSKKQVC-----SVDSSSAEWSKLNGNFMNGPSENL-----AGK--MNDQASVE 360  
282 VPT-----AERDLSGKGVSD--LSLSGSVC-----CKHKPSSTSLSP 319  
361 LTVNTSLARNSDSSSEIKVINATILSSEKSPCVDFIRGSGLSCHYSILSEVQYSTC 420  
320 RQ-----NA-----POLPPIETRHFLYSRGGQGERLNPFLIS-N 354  
421 PQVGPSSYLHNSCFCFNCIISNASKHLSLDSLSIRNGIFYNRTTYSVPHCHILSKR 480  
355 LQPNLTGARRLVEIIP-----LGSRRPTSGPLCRTHLSRRRWQMPFLFOQ 400  
481 KRPD--ALSUVKHIFGINSCCASLLKYNCHSTIRKSNCLC-----CWLPSKI-KN 528  
401 LLYNHAEQVRLRLRHCHFR-----TANQVTDALNTSPP-----HLMDDLRLH 445  
529 LIENSKRCQYKFLFKHCSVKCKAPDVTKN-----DGKAHYPPGGAAYDRSPSRL 584  
446 SSPQVQVGLRACLCKVVASLWGTHTHNERFPFNKLLKFI SLGKYGKLSLQELMWKQVE 505  
585 STHQQVASFVAVLRKIVPKPLNGFPFGKSLRUTNIWPKILRPFQPSDCDGLDKVS 644  
506 DCHWLRSPP-----GKD--RVPAAEHLRLERILATEFLWIMDTYVQLLSRSPFI 553  
645 HYSWLNIENFSCFCGAIIGKQTGSTSAEEQEKVILLHCWISLMSFLDIVIPVYRTFYV 704  
554 TESTFOKNLRFVRKSVWKLQSIGVRQHLEVRRLRSQ-----EVRHQDTWAMPICR 610  
705 TERESKRYDVFIYPKSVWDLTNSAIAS-LNKKCNFRILRGEPKRAVRLNCS-----SR 757

```
57 CMHWSQPP-----PADLSFHOVSLSKELVARVQVORLCERNERNVLAFFGELLNEARGSP 111
90 AADDLSAPPPPPPTPGNSGHSQALVEMKMSIVADO-SHGTKNVLCNGUH-----EGGQ 143
112 PMAFTSSVRSYLPNTVITETLRVSGAMWLLSRVGGDLLVYLLAHALYLLVPPSCAYQVC 171
144 SIC-----ISDLVSSSSWSILLHRIGDLMLCYLLCTSIIFLPVKNDYFQVS 190
172 GSPLYQI-----CATTDIWS-----VSASY 192
191 GVLNVLRNFIFASTVARKHQPTTKAKCHCTCYLWKSANMAENLSICHDSNSGVNSSF 250
193 RPTREV--GRNTNLRFLQOIKSSROEAPKPLALPSRG-----TKRHLSLTS 238
251 SSTKIVTQOCETCGSIRRAESKDPGECNCPKPSDGRSGCCNCTHNTKRKRLYS 310
239 TSVSAKACACYPVPRVE-----GPHRQVLPTPSGSKWSPSPARSP 281
311 WORSKKQVC-----SVDESSAEMSKLNGFNFMNSGPFSENL-----AGK--MNDQAQSYE 360
282 VPT-----AEKDLSSKGKYSV--LSLSGSVC-----CKKPSSTLSLSP 319
361 LTVDNLSLARSNDSSSEIKVINATILSSEKSPCSVFDIRGSOGLSCHYSLSVEVQVQSTC 420
320 RQ-----NA-----FOLRPIETRFYLSRGDQGRNLNPSFLLS-N 354
421 PQVGESSYLHNSCICFCNCIISNASKHLSLDSLSIRNGIFYNRRITTSYFHCXHLISKR 480
355 LOPNLGTARRVLEIF-----LGSRPSTGPLCRLTHLSRRYQMRFLPQ 400
481 KPDP-----ALSIVKHIFGINSCCASLLKYNCHESITKSNCLC-----CWLPKSI-KN 528
401 LLVNHAECQYVRLRSHCRFR-----TANQQVTDALNTSP-----HLMDLRLRH 445
529 LIRNSKRQYKXKFLKHGCVKCKVAPVTXN-----DGKAHYPPGKAAAYDRSFRLEAY 584
446 SSPWQVYGLACLVKCKVYSASLMTGRNERFFNOLAKKFLSLGKYKLSLOELMWKQYE 505
585 STHQQVASFVAVLKRIVPKPLNGSGFKSLRNTIWKFLKRRFETQLSDCIGDLKVS 644
506 DCHWLRSSP-----GKD--RVFAAEHRLRERILATFLWLMDTYVVQLRFFFYI 553
645 HYSWLSNIEFNSFCFSAIIGKTGSTSAEQKQNLHWCMLPSDVIPIVVRTYFVY 704
554 TESTQKRLFPYKSVKSLQSGIVGQHLRVLRELSQL-----EVRHQDTWLAMPICR 610
705 TERESKRYDVFPYKSVWRDLTSNAIAS-LNKKCNFRILRGEPRKAVRLNCS-----SR 757
611 LRFIPKPNGLRPIVNMYSYMGTRALGRRKQAOHFTQRLKTLFSLMNLVYRTKPHLMSGSV 670
758 VFPLPKADMRPLVDL-----RAKSDANLKNCHILMKLXDEKPEMGSSV 804
671 LGMNDIYTRAFVLRVA--LDQTPRYFVKADVTGAYDAIPGKLVEVVANMIRHSEST 729
805 FDNVNHQNLQFISSEKASQLMKKLKVVIVVADVSKAFDCVSHDWMLKMDIDDAFKODE-- 862
730 YCIRQYA--VVRDSOGVHKSFRRQVITSLDLPYMGQFLKHLDQSDASALRNSVVEIOS 788
863 YTVKCKSVKVCNRSKSLYF-----DSNAGNSGNSTYDLS 899
789 ISMNESSSLF-----DFFLHFL-----RHSVYKIGDRCYTQCQIGPQSSLSLT 833
900 IQLS-SGGGIFVDQGTICRLKEQFPHLLVBOIKCNILKIGQYKYLQVQVIAQSKLSPN 958
834 LCSLCFGDMENKLAERVQDVG-----LLLRFDVDDLLVTPHLDQAKTFLSTL 880
959 LCSLYGHLENSVSKFUDHDSKLNAGEAFSPFYLMLRFIDDFIFISFSLHFAQKFLNRM 1018
881 VHGVEPYGCMNLQKTVTVNFFVEPQTGGAPYQLPAH-----CLFPWCGLLLDQTLE 934
1019 REGFVFNKYNMDSKYGNF-----CAGNSEPSSNRLYRGDGDGVSFNPWSGLLINCETLE 1073
```

```
935 VFCDYSGYACTS--IKTSLTFSQVFKAGTKMTNKLKSLVRLK-CHGLELDQVNSLQTV 991
1074 IQADYTRYLDITHYLPSSLVRLCHLQHSIST--SKLCLPYAVQNAHPFYDSNINSPTGIR 1131
992 INYKIFLLQAVRFHACVITQLPFDQVRK 1020
1132 VNITYQAFLLCAMKEH-CYIRSVSDANVSK 1159

RESULT 14
Q9JULM1 PRELIMINARY; PRT; 104 AA.
AC Q9JULM1;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Telomerase reverse transcriptase (Fragment).
GN TERT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hatakeyama S., Ishikawa F.;
RT "Identification of the mouse telomerase reverse transcriptase (mTERT)
promoter.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF157502; AAP42984.1; --
DR MGD; MGI:1202709; TERT.
KW RNA-directed DNA polymerase.
FT NON_TER 104 104
SQ SEQUENCE 104 AA; 11998 MW; 4B649B63476D3D44 CRC64;

Query Match 9.3%; Score 550; DB 11; Length 104;
Best Local Similarity 100.0%; Pred. No. 9.9e-39;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRAPCPAIVRSYLLRSRYREVWPLATFVRLPGEGRRLVQCPDKPIYETLVAQCILVCMHW 60
DB 1 MTRAPCPAIVRSYLLRSRYREVWPLATFVRLPGEGRRLVQCPDKPIYETLVAQCILVCMHW 60
QY 61 GSOPPPADLSFHQVSSLSKELVARVQVORLCERNERNVLAFFGELL 104
DB 61 GSOPPPADLSFHQVSSLSKELVARVQVORLCERNERNVLAFFGELL 104

RESULT 15
Q8MUB3 PRELIMINARY; PRT; 1032 AA.
AC Q8MUB3;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Telomerase reverse transcriptase.
GN TERT-1.
OS Euplotes crassus.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
OC Euplotida; Euplotidae; Moneuplotes.
OX NCBI_TaxID=5936;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L., Dean S.R., Shippen D.E.;
RT "Oligomerization of the Telomerase Reverse Transcriptase from Euplotes
crassus.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF528527; AM95622.1; --
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR003545; Telomerase_RT.
DR Pfam; PF00078; rvt; 2.
DR PRINTS; PR01365; TELOMERASERT.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1032 AA; 122966 MW; 0C401515839801A8 CRC64;
```

Query Match 7.6%; Score 450; DB 5; Length 1032;  
Best Local Similarity 22.7%; Pred No. 8.6e-29;  
Matches 179; Conservative 142; Mismatches 316; Indels 150; Gaps 32;  
392 WQRPPLFQQLLVNH-----AECQYVLLRSHCRPTANQOQVTDALN-----TSPPHLMD 440  
323 WVER---KKVILSHLNDRAKAKAYEELFS---YTIENKCVTOFLNEFFYHTLPKDFM- 375  
441 LLRLHSPWQVYGFACLCKVVSASLWTRNERRFFKNLKPISLGKYGKLSLQELMW 500  
376 -----TGKRRKSPQSKIKKYVGLNKHLELHKNLIG 406  
501 KWKVEDCHWLRRSPGK-----DRVPAEHRRLERILATFLWMDTYVQVLLRSFYIT 554  
407 KLDIKSIKWLGFKTSKKNHYFDK-----ENRF---VLWRILRWIPEDVVVSLIRCFYVT 459  
555 ESTPQKURLFFYKSVMSKQSIGVROHLERVRLRELSEEVRRHODTWLAMPICRLRFI 614  
460 EQKSYSETYYIKNIWDIVMKYSIAD-LNRETLEEVPEKEDWVKWKNELRFAP-GKLRLI 517  
615 PKPGLRPIYVMSYMGTRALGRRKQAQHTQRLKTLFSLMYERTKPHL---MGSSVL 671  
518 PKTTFAPIN---TFNKKILDPEGKTSQMTTNTKLSYSHLMLKTLKRMFKOPFGPAVF 573  
672 GMDIVETWEAFVLRVRLDQTPMYVFKADVTGAYDAIPOQKLEVVVYANMIRHSESTYC 731  
574 NYDDWAKYEFVLRQVGR-FGLYFVTWIDEKYDSVDREKLSQLLGT-TRLSEFR 631  
732 IRQYAVVRD-----SQQVHKSPR---ROVTTLSLQPYMGQFLXHQD--SDA 776  
632 IMTVQAMKRNNAEVDLDKCAKNKKECFQKQKIALEGDQYP---SLNVLEDQNDL 688  
777 SALRNSVVIQSI SMNESSSLEDFLHFLRHSVVKIGRCVTCOCGIPQGSSTLLCS 836  
689 NA-KETLLVENQRDPYKXALLDPVIKICRNYIEFNRYKYQKPGPQGLCVSSILSS 747  
837 LCFQDMENKLF AEVQRD-----LLRFVDDPFLVTPHLDQAKTFLSTLVHGVPEY 887  
748 FYASLEENALGYLRKESMDANDPNTLLMLRLDDYLLITTKENNAILFIEKLINVRQN 807  
888 GCMINLOKTVNPFVPGTLG---GAAPYOLPAHCLFPWCGLLDQTLF----- 934  
808 KFKENMKKJQTNFPLDPSKLNKYGMASVEDQNIADHYIDWIGISIDMTALMNPINLR 867  
935 --VFCDYSGYAQTSIKTSLTFQSVFVFRAGTKMKNKLSVLRLKCHGLFLDLQV---SLQT 989  
868 KGILCTLNMNQTK-----KASWMLKRLKSLFLMNNITHYPRKTTITNREFSNKT 916  
990 VCINIIYKIFLLQAYRPHACVQLPPDQVRKNL-----TFFGLISS--QASCCYAIL 1040  
917 ----LNKLYTAGAYKNQCCIE--YKDHFKTNTTEHPQLDKIICAIYISVTRAFKYLVC 970  
1041 KVKNPGMTLKASGSFP-PEAAHWLCYQAFLLKLAHSVIYKICLLGLPLRTAQKILCKLPE 1099  
971 NIKDSIYDREHSRDFSTATLRHI--EIFSLKKNHFGVVKILVA---KEQKMLEEK-DN 1024  
1100 ATWTLK 1106  
1025 ACMTAIE 1031

rch completed: February 4, 2004, 14:11:48  
, time : 50 secs